

High-throughput sequencing reveals inbreeding depression in a natural population

Joseph I. Hoffman^{a,1}, Fraser Simpson^b, Patrice David^c, Jolianne M. Rijks^d, Thijs Kuiken^e, Michael A. S. Thorne^f, Robert C. Lacy^g, and Kanchon K. Dasmahapatra^h

^aDepartment of Animal Behaviour, Bielefeld University, 33501 Bielefeld, Germany; ^bDepartment of Genetics, Evolution, and Environment, University College London, London WC1E 6BT, United Kingdom; ^cCentre d'Ecologie Fonctionnelle et Evolutive, Centre National de la Recherche Scientifique, 34293 Montpellier, France; ^dDutch Wildlife Health Centre, Pathology Division, Department of Pathobiology, Faculty of Veterinary Medicine, University of Utrecht, Yalelaan 1, 3584 CL, Utrecht, The Netherlands; ^eDepartment of Viroscience, Erasmus University Medical Centre, 3000 CA, Rotterdam, The Netherlands; ^fBritish Antarctic Survey, Natural Environment Research Council, Cambridge CB3 0ET, United Kingdom; ^gDepartment of Conservation Science, Chicago Zoological Society, Brookfield, IL 60513; and ^hDepartment of Biology, University of York, Heslington, York, YO10 5DD, United Kingdom

Edited by Brian Charlesworth, University of Edinburgh, Edinburgh, United Kingdom, and approved February 6, 2014 (received for review October 21, 2013)

Proxy measures of genome-wide heterozygosity based on approximately 10 microsatellites have been used to uncover heterozygosity fitness correlations (HFCs) for a wealth of important fitness traits in natural populations. However, effect sizes are typically very small and the underlying mechanisms remain contentious, as a handful of markers usually provides little power to detect inbreeding. We therefore used restriction site associated DNA (RAD) sequencing to accurately estimate genome-wide heterozygosity, an approach transferrable to any organism. As a proof of concept, we first RAD sequenced oldfield mice (*Peromyscus polionotus*) from a known pedigree, finding strong concordance between the inbreeding coefficient and heterozygosity measured at 13,198 single-nucleotide polymorphisms (SNPs). When applied to a natural population of harbor seals (*Phoca vitulina*), a weak HFC for parasite infection based on 27 microsatellites strengthened considerably with 14,585 SNPs, the deviance explained by heterozygosity increasing almost fivefold to a remarkable 49%. These findings arguably provide the strongest evidence to date of an HFC being due to inbreeding depression in a natural population lacking a pedigree. They also suggest that under some circumstances heterozygosity may explain far more variation in fitness than previously envisaged.

inbreeding | genetic variability | heterozygosity fitness correlation | single nucleotide polymorphism | RAD sequencing

It has long been known that inbreeding can reduce fitness, mainly through the unmasking of recessive or partially recessive deleterious alleles (1). Such effects are well documented in the laboratory (2, 3) but until recently have remained largely unstudied in natural populations (4, 5). This is because an individual's inbreeding coefficient (f) can be calculated directly only by using a deep pedigree, and these pedigrees are seldom available outside the laboratory (6). However, because inbreeding increases an individual's homozygosity, the heterozygosity of a panel of neutral genetic markers can in theory be used as a surrogate for f .

Initial searches for heterozygosity fitness correlations (HFCs) (7, 8) used allozymes, but the results obtained are difficult to interpret because the proteins themselves may be under selection. With the discovery of microsatellites, an abundant class of putatively neutral genetic marker, it became possible to look for HFCs without concerns about selection on the markers themselves. A burgeoning literature now shows that HFCs based on small numbers of microsatellite loci are found in many bird and mammal species for a remarkable range of fitness traits, including neonatal survival (9), parasite susceptibility (10) and lifetime reproductive success (11), and even behavioral qualities such as territory-holding ability (12), song complexity (13) and attractiveness (14). This weight of evidence suggests that HFCs are an important and widespread phenomenon in the natural world. It is therefore important to understand their basis.

Theory predicts that HFCs arise as a result of inbreeding depression, which will reduce the fitness of individuals in pro-

portion to their inbreeding coefficient f (7). Therefore, variance in inbreeding coefficients within a population is necessary to generate HFCs. However, simulation and empirical studies indicate that the variance in f in natural populations is usually low, and estimates of f based on the small numbers of markers typically deployed are often very poor (15). As a result, HFCs will usually be weak or nonsignificant even when inbreeding actually explains a large proportion of fitness. Szulkin et al. (16) summarized this by saying that HFCs allow one to see only the “tip of the iceberg” and provided two examples in which heterozygosity explains 3% and 6% of trait variation even though inbreeding is expected to account for 24% and 30% of the variance, respectively. The most favorable situations to observe HFCs are in populations where inbred individuals are not rare due to factors such as small population size, extreme reproductive skew, and natal philopatry (15), but even in these cases using only a few markers severely curtails the power to observe the actual impact of inbreeding.

To explain HFCs, authors also frequently invoke “local effects” where one or a small number of the microsatellites used as markers are by chance linked to a gene experiencing heterozygote advantage (8). Local effects are widely discussed in the literature, but their importance is unclear given that balanced polymorphisms are

Significance

Many studies of wild populations reveal links between heterozygosity and fitness, with relatively heterozygous individuals carrying fewer parasites, living longer and being more attractive to mates. These patterns appear ubiquitous and are often highly significant, but heterozygosity usually accounts for very little of the total variation in fitness. However, most studies analyze only around 10 loci, representing a tiny fraction of the genome. We therefore used high-throughput DNA sequencing to estimate genome-wide heterozygosity based on over 10,000 loci and found it to accurately reflect inbreeding. Applied to wild harbor seals, genome-wide heterozygosity explained almost half of the variation in parasite infection. By implication, a greater proportion of fitness variation could be linked to genotype than previously thought.

Author contributions: J.I.H. and K.K.D. designed research; J.I.H., F.S., J.M.R., T.K., R.C.L., and K.K.D. performed research; J.I.H. and K.K.D. contributed new reagents/analytic tools; J.I.H., P.D., M.A.S.T., and K.K.D. analyzed data; and J.I.H., P.D., and K.K.D. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Data deposition: The sequences reported in this paper have been deposited in the Short Read Archive database (accession no. [PRJEB5164](https://www.ncbi.nlm.nih.gov/sra/PRJEB5164)).

¹To whom correspondence should be addressed. E-mail: joseph.hoffman@uni-bielefeld.de.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1318945111/-DCSupplemental.

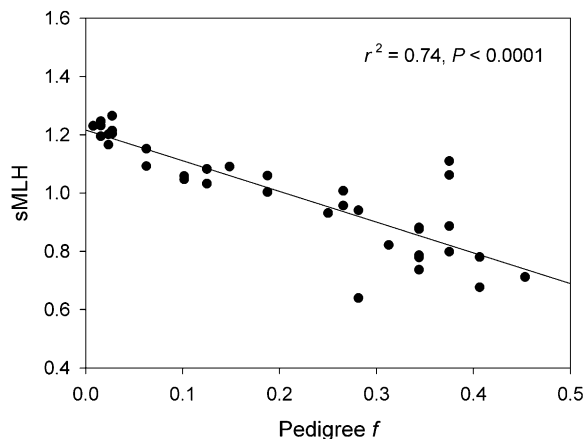


Fig. 1. Correlation between pedigree-based inbreeding coefficient (f) and RAD standardized multilocus heterozygosity (sMLH) for 36 oldfield mouse individuals. The r^2 for the equivalent relationship based on 12 microsatellites is substantially lower at 0.276.

thought to be rare (17) and strong linkage between random pairs of loci is infrequent (16). It has also been argued that the contribution of local effects to HFCs may be overstated due to many studies having used an inappropriate statistical framework (16).

One means of unambiguously testing whether inbreeding alone can explain HFCs is to deploy a larger number of genetic markers (18). If an HFC is due to inbreeding depression, the use of more markers will reduce the error variance in the estimation of genome-wide heterozygosity and thereby strengthen the correlation. In contrast, if the HFC is highly dependent on one or a few marker loci being by chance linked to a fitness locus with strong effects, as in the local effect model, adding more markers located throughout the genome will reduce the strength of the HFC as the contribution of any one marker becomes progressively diluted. Until recently, however, this approach was not available to most studies because of the prohibitive costs of developing and screening large numbers of additional microsatellites.

Restriction site associated DNA (RAD) sequencing (19) has recently emerged as a rapid and economic means of genotyping thousands of single nucleotide polymorphisms (SNPs) in virtually any organism. This approach concentrates high-throughput sequencing effort around restriction enzyme cut sites that are distributed across the genome, thereby generating sufficiently deep local sequence coverage to reliably call SNPs as being either heterozygous or homozygous. Although SNPs are individually less informative than microsatellites due to their lower allelic diversity, this should be more than compensated for by the large numbers of markers screened, in principle allowing genome-wide heterozygosity to be estimated with far greater precision than previously possible.

Here, we evaluated the ability of RAD sequencing to accurately estimate genome-wide heterozygosity using an experimental population of oldfield mice (*Peromyscus polionotus subgriseus*) with known pedigree-based inbreeding coefficients (20). We then applied RAD sequencing to a natural population of harbor seals (*Phoca vitulina*) to determine whether inbreeding explains previously reported HFCs for survivorship and parasite infection (21). We hypothesized that these HFCs should strengthen with the deployment of many thousands of SNPs if they are due to inbreeding.

Results

Relationship Between RAD Heterozygosity and f . Illumina RAD sequencing of 36 oldfield mice yielded 265 million paired-end reads (Fig. S1) that were assembled into 63,129 RAD tags (Fig. S2A) as

described in *SI Methods* and *SI Results*. After SNP calling and filtering (Figs. S3 and S4), 13,198 tags were retained for further analysis, each of which contained a single biallelic SNP (Fig. S5). Individual heterozygosity was expressed using the measure sMLH, which corrects for the fact that not all individuals are genotyped for the same loci by standardizing average multilocus heterozygosity in an individual by the average observed heterozygosities in the population of the subset of loci for which the focal individual was genotyped (10). RAD sMLH was strongly correlated with pedigree-based f (Fig. 1, $r^2 = 0.74$, $P < 0.0001$), far more so than sMLH at 12 microsatellites ($r^2 = 0.28$, $P = 0.0003$). A highly significant relationship was also obtained between RAD allele sharing and pedigree-based pairwise relatedness ($r = 0.94$, $P < 0.0001$), which was stronger than the equivalent relationship based on microsatellites ($r = 0.68$, $P < 0.0001$).

Variation in inbreeding within a population increases the variance in sMLH beyond that expected if heterozygosities at different loci were statistically independent. As expected, this genetic signature of inbreeding is conspicuous in a plot of the distribution of RAD sMLH in our sample of oldfield mice (Fig. 2A). The excess in variance is quantified by a parameter called g_2 , which measures the extent to which heterozygosities are correlated across loci and can be computed from genotype data using the method of David et al. (22). In our sample of oldfield mice, g_2 calculated from 13,198 SNPs was 0.035 ± 0.007 SD, which is significantly different from

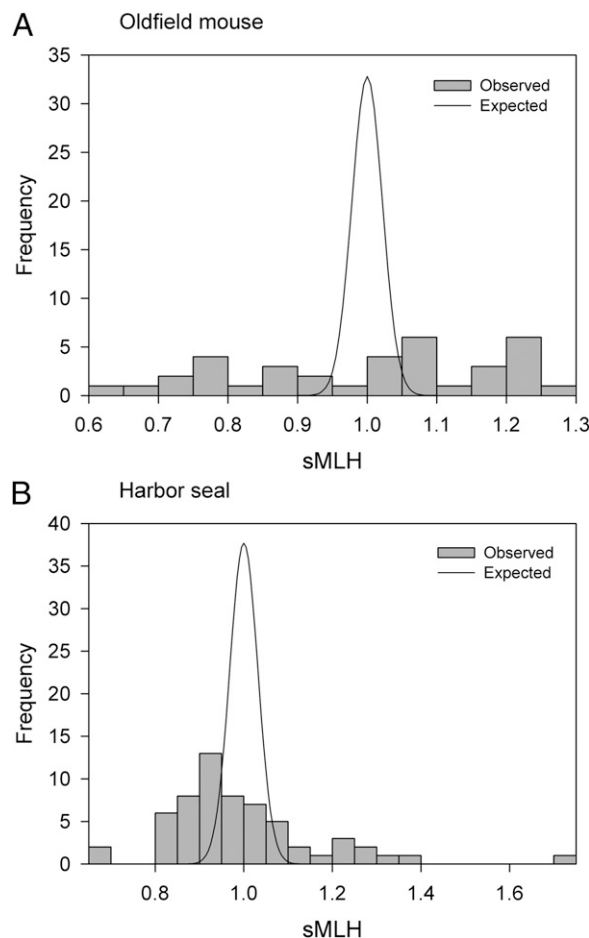


Fig. 2. Observed distributions of RAD sMLH (shaded bars) together with expected distributions under the null hypothesis of $g_2 = 0$ (solid curves), meaning that there is no variance in inbreeding. See *Methods* for details. Results are shown separately for (A) oldfield mice genotyped at 13,198 SNPs and (B) harbor seals genotyped at 14,585 SNPs.

zero ($P < 0.001$). With so many markers, the sampling of loci does not result in any appreciable estimation error, and even with as few as 2,000 SNPs this error is still very small (Fig. S6).

The large number of SNPs deployed also suggests that RAD heterozygosity should provide a very precise estimate of individual f : From table 2 in Szulkin et al. (16) the expected correlation between sMLH and f , denoted $r_{sMLH,f}$, is $-1.03 (\pm 0.024 \text{ SD})$. Pedigree-based estimates of f can be used to provide direct estimates of g_2 and of $r_{sMLH,f}$, which turn out to be slightly lower ($g_2 = 0.028 \pm 0.009 \text{ SD}$; $r_{sMLH,f} = -0.86$) than those based on the RAD data. As discussed later, this is to be expected because pedigree f does not capture all of the variance in inbreeding due to unknown founder relatedness, possible pedigree errors, and the random deviation of actual heterozygosity from the statistical expectation.

The accuracy of RAD-based heterozygosity estimates was further confirmed by dividing the markers into two random subsets, computing the correlation in sMLH between the two, and repeating this 1,000 times to obtain a distribution of heterozygosity-heterozygosity (het-het) correlation coefficients (15). The resulting values are tightly centered on a mean of $0.975 (\pm 0.006 \text{ SD}$; Fig. 3) for the RAD data, whereas the equivalent value based on 12 microsatellites is only $0.187 (\pm 0.125 \text{ SD})$. Again, as few as 2,000 SNPs are sufficient to obtain a strong positive correlation (Fig. S7A). This suggests that genome-wide heterozygosity can be reliably estimated from the many thousands of markers generated by RAD sequencing.

Data on body mass at weaning were also available for the oldfield mice, allowing us to quantify the strength of association with f and marker heterozygosity. Separate generalized linear models (GLMs) of mass were constructed in which f , microsatellite sMLH, and RAD sMLH were fitted respectively as predictor variables, together with sex and the two-way interactions. Only the genetic terms were retained in the reduced models, in which f was significant at $P = 0.027$ ($F_{1,35} = 5.35$), microsatellite sMLH at $P = 0.037$ ($F_{1,35} = 4.69$), and RAD sMLH at $P = 0.025$ ($F_{1,35} = 5.51$). Thus, RAD sMLH is at least as good a predictor of body mass as pedigree f .

Application to a Natural Vertebrate Population. Sixty harbor seals were RAD sequenced, generating 374 million paired-end reads (Fig. S1). These were assembled into 83,148 RAD tags

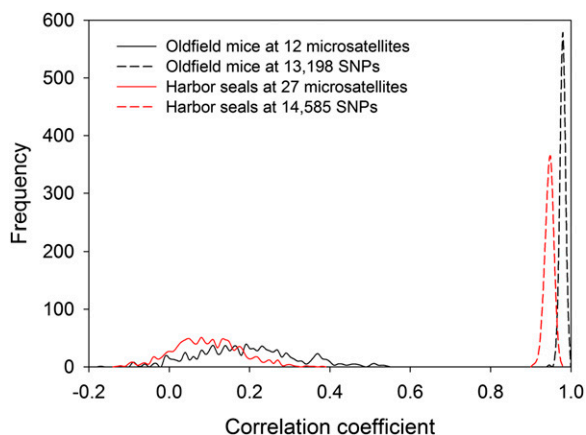


Fig. 3. The strength of correlation in heterozygosity among loci. The markers for which animals were genotyped were randomly divided into two equal groups and the correlation coefficient (r) between the resulting sMLH values was calculated. This procedure was repeated 1,000 times to generate distributions of r values. Values centering around zero suggest that inbred individuals are rare or absent, whereas increasingly positive values indicate that inbred individuals are present.

(Fig. S2B), of which 14,585 were retained for further analysis (details in *SI Methods* and *SI Results*). Similarly to the oldfield mouse, a strong signal of inbreeding was present in the RAD data, with g_2 estimated at $0.028 \pm 0.009 \text{ SD}$ ($P < 0.001$), which in our experience is a high value for a natural population (16). This is also illustrated by the empirical sMLH distribution (Fig. 2B), which has a much larger variance than expected in the absence of variation in inbreeding. Consequently, RAD heterozygosity is a very precise estimate of individual inbreeding ($r_{sMLH,f} = -0.97$) in this sample, whereas the equivalent number based on 27 microsatellites is only -0.21 . Accordingly, the mean het-het correlation coefficient increases from $0.090 \pm 0.084 \text{ SD}$ based on 27 microsatellites to $0.941 \pm 0.011 \text{ SD}$ (Fig. 3) for the RAD data. This signal is slightly weaker than in the oldfield mouse both overall and when equivalent-sized random subsets of SNPs are considered (Fig. S7A). However, this is expected, as g_2 is lower than in the experimental oldfield mouse population where inbreeding coefficients range from 0 to 0.45. Moreover, the SNPs are on average less heterozygous in the harbor seals than in the mice (mean unstandardized MLH = $0.062 \pm 0.011 \text{ SD}$ and $0.164 \pm 0.029 \text{ SD}$, respectively) and are therefore relatively less informative (Fig. S5).

If an HFC is due to inbreeding depression, adding more markers should improve the estimate of f , and the relationship between fitness and average marker heterozygosity should strengthen (18). To test this prediction empirically, we fitted RAD sMLH as a predictor variable in GLMs of two fitness-related traits, longevity and parasite infection, both of which were coded as binary response variables (details in *SI Methods*). Young harbor seals, defined as being less than 1 y of age, showed a nonsignificant tendency to be less heterozygous at microsatellites than older seals (Fig. 4, $F_{1,59} = 2.09$, $P = 0.15$). This difference strengthened to become marginally significant with RAD heterozygosity ($F_{1,59} = 4.39$, $P = 0.04$). However, when a single highly heterozygous outlier (RAD sMLH = 1.73, two-tailed Grubbs' test, $\alpha = 0.05$) was removed from the analysis, RAD heterozygosity no longer remained significant ($F_{1,58} = 3.14$, $P = 0.08$).

As infection and mortality due to lungworm burden occur mainly during the first year of life, we restricted our analysis of parasite infection to the young seals, 23 of which were infected with lungworms. We found that a close to significant trend for lungworm-infected seals to have lower microsatellite heterozygosity than uninfected seals (Fig. 4, $F_{1,29} = 3.51$, $P = 0.06$) became highly significant with the RAD data ($F_{1,29} = 16.03$, $P = 6.23 \times 10^{-5}$). Moreover, the deviance in lungworm infection explained by heterozygosity increased almost fivefold from 10.8% to 49.2%. By implication, genome-wide heterozygosity strongly influences susceptibility to lungworm infection.

To further test the prediction that an HFC should strengthen with marker number if inbreeding depression is responsible, we reran the GLM of parasite infection after randomly sampling five subsets each of between 50 and 14,000 SNPs and recalculating RAD sMLH. The average percentage deviance explained by heterozygosity increased gradually with SNP number before leveling off at around 7,000 loci (Fig. S7B). This pattern is very similar to the observed relationship between the number of randomly subsampled SNPs and the mean correlation in sMLH among markers (Fig. S7A). We also explored sensitivity to SNP minor allele frequency (MAF) by partitioning the marker set into loci with MAFs above and below 0.1, recalculating RAD sMLH for both and fitting these together in the model. Both categories of SNP were highly significant and thereby contribute toward the overall signal (MAF < 0.1 , $F_{1,29} = 18.40$, $P = 1.79 \times 10^{-5}$; MAF > 0.1 , $F_{1,29} = 8.27$, $P = 0.004$). That the lower-frequency SNPs were more significant probably reflects the far larger sample size of loci (12,455 vs. 2,130, respectively).

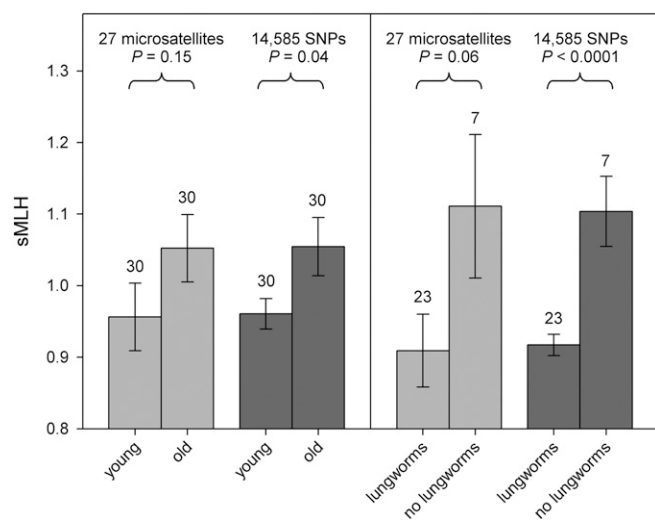


Fig. 4. Relationship between sMLH, estimated from microsatellites and RAD genotypes, and two fitness-related traits in harbor seals. Bars with light shading represent mean (\pm SEM) microsatellite sMLH and bars with dark shading represent mean (\pm SEM) RAD sMLH. Sample sizes are given above the bars. (Left) The mean (\pm SEM) sMLH of young and older seals, the former being classified as those that died before reaching 1 y of age (details in *Methods*). A nonsignificant tendency based on 27 microsatellites for younger seals to have lower heterozygosity than older seals strengthens to $P = 0.04$ when 14,585 SNPs are deployed. (Right) The mean (\pm SEM) sMLH of young seals with and without lungworms. A nonsignificant trend for infected seals to have lower heterozygosity than uninfected seals based on 27 microsatellites becomes significant at $P < 0.0001$ with 14,585 SNPs.

Discussion

We evaluated the ability of RAD sequencing to estimate genome-wide heterozygosity using a pedigreed oldfield mouse population with substantial inbreeding (up to $f = 0.45$) over six generations. We show that RAD heterozygosity is far more strongly correlated with f than microsatellite heterozygosity and that the r^2 further increases to 0.93 if three outliers are removed. Additionally, the RAD data yield high het-het correlation coefficients and RAD sMLH is a marginally better predictor of body mass than f . Taken together, these findings suggest that RAD heterozygosity is a better estimate of genome-wide heterozygosity than pedigree-based estimates of f in our study. Although the exact reasons for the three outliers are unclear, in this case they are more likely the result of pedigree errors or mislabeled samples than unknown levels of relatedness among the founders. First, any effects of founder relatedness would be propagated throughout the pedigree and should therefore affect the majority of individuals. Second, although we do not have RAD genotypes of the founders, we were able to obtain relatively crude estimates of founder relatedness based on microsatellite and amplified fragment length polymorphism data reported previously (20). Adjusting the pedigree inbreeding coefficients with these estimates of founder relatedness did not improve the correlation with RAD heterozygosity. Regardless of the exact causes, our findings suggest that typing a very large number of loci may under certain circumstances allow one to estimate inbreeding even more precisely than from a pedigree. Furthermore, RAD sequencing provides a tool for error checking and otherwise refining pedigrees.

SNPs are attractive markers for correlating genome-wide heterozygosity with fitness (23). Although they are individually less variable than microsatellites, SNPs can be rapidly and economically genotyped in large numbers. Even so, it has been argued that for a large mammalian genome, a “herculean survey of 3,000 markers” will be required to produce a correlation between

heterozygosity and f of around 0.4 (24). This is supported by two recent studies that correlated sMLH with f in captive zebra finches, obtaining r values of -0.07 based on 771 SNPs (25) and -0.46 based on 1,359 SNPs (26). However, we show that RAD sequencing is capable of screening the heterozygosity of at least an order of magnitude more SNPs, bringing a commensurate improvement in the estimation of genome-wide heterozygosity. In our study, much of the inbreeding signal is captured by a couple of thousand markers (Figs. S6 and S7), but clearly the number of loci required in other systems will depend upon the extent to which genome-wide heterozygosity varies among individuals (15). This is currently unknown for most natural populations, although approaches such as ours make it increasingly amenable to study.

A number of factors could potentially introduce bias into our sMLH estimates, although these would be expected to affect most individuals within a sample to a similar extent. For example, in the absence of linkage maps we were unable to discriminate between SNPs on autosomes and sex chromosomes, meaning that our estimate of heterozygosity for males will be slightly lower than heterozygosity averaged over the autosomes. A second bias stems from the fact that RAD markers are affected by nonamplifying “null alleles,” which result mainly from mutations in the restriction enzyme recognition site. Based on the number of RAD tags without any SNP polymorphisms in our datasets, the level of nucleotide polymorphism, π , is estimated to be of the order of 0.002. Therefore, with an 8-base restriction enzyme, null alleles are expected to be found in at most only 2% of RAD markers. Null alleles are an issue for many marker types, including microsatellites, and result in a downward bias in estimated heterozygosity. Third, RAD markers paired on either side of restriction cut sites are in strong linkage and are therefore nonindependent. In our datasets, as only 10–20% of RAD tags contain a SNP (*SI Results*), the likelihood of tags on either side of a restriction cut site each carrying a SNP is relatively low. Consistent with this, based on calculations of pairwise linkage disequilibrium among all SNPs, $\sim 10\%$ of SNPs are estimated to be in strong linkage disequilibrium with another SNP in both study systems. These factors will introduce some imprecision into estimates of heterozygosity and may be responsible for some of the unexplained variance in Fig. 1.

Despite these caveats, there are reasons to believe that RAD sequencing may produce less biased estimates of heterozygosity than many other SNP genotyping approaches. High-density genotyping arrays, for example, are usually based on SNPs identified from small “discovery panels” of individuals. This can distort the allele frequency spectrum, as SNPs with intermediate allele frequencies tend to be overrepresented and rare SNPs are often missing (27). As a result, estimates of heterozygosity based on the polymorphic SNPs tend to be inflated, whereas genome-wide heterozygosity will be underestimated because low-frequency SNPs are ignored (28). Thus, although no single approach is perfect, RAD sequencing at least eliminates ascertainment bias because the entire sample serves as the discovery panel.

Contrasting MAF distributions were obtained for the oldfield mice and the harbor seals, the latter being strongly left skewed (Fig. S5). Although we are unable to discount the possibility that at least some of the low-frequency variants in the harbor seals could be technical artifacts, we thoroughly explored the SNP calling parameter space in the oldfield mouse, using pedigree information to derive optimal parameters that were then applied equally to both species (*SI Methods* and *SI Results*). We also guarded against spurious genotypes by constructing the reference genome against which the reads were aligned using tags that were present in at least two individuals. Finally, SNPs with MAFs below 0.1 were found to contribute significantly toward the overall HFC signal, suggesting that they carry useful information about inbreeding. A plausible explanation for these contrasting distributions is therefore that they reflect species-specific differences in population size and history. Oldfield mice occur

across much of the southeastern United States and can reach densities of 18–26 animals per hectare (29). Relatively high average heterozygosity and an abundance of SNPs with intermediate allele frequencies could therefore be explained by a large and stable effective population size. In contrast, many pinniped species show evidence of strong demographic responses to historical changes in the marine environment, usually through rapid population expansion during periods of increased food or habitat availability (30, 31). Little is known about the historical demography of harbor seals, although the Waddensee population has been reported to have low heterozygosity (32), which is consistent with the effective population size having at some point in the past been very small. The excess of low-frequency polymorphisms could therefore be attributable to postglacial population expansion.

Het-het correlations (15) are often used to test for the presence of inbred individuals, but with only around 10 markers commonly deployed, this approach lacks power because the marker subsets are very small. This is exactly what we found (Fig. 3): the distribution of microsatellite-based het-het correlation coefficients in the oldfield mouse was slightly positive but with a large SD, despite the sample containing highly inbred individuals ($f \leq 0.45$). In contrast, the RAD data yielded het-het correlation coefficients well in excess of 0.9 for both the oldfield mice and harbor seals. This provides clear evidence for inbred individuals being present not only in the mouse pedigree but also in the natural harbor seal population.

A second line of evidence also points toward the HFCs for survival and parasite infection being due to inbreeding depression in harbor seals. With inbreeding, the deployment of additional markers should strengthen the HFC detected, whereas with a local effect adding more markers is expected to weaken the relationship between mean heterozygosity and fitness (18). We found that the microsatellite-based HFC for lungworm infection strengthened considerably with RAD heterozygosity, the statistical significance increasing from $P = 0.06$ to $P < 0.0001$. Such a dramatic strengthening of effect has not been previously reported, although one study was able to replicate an HFC for birth weight (but not one for juvenile survival) using independent panels of 10 microsatellites (33). Furthermore, the deviance in lungworm infection explained by heterozygosity steadily increases with the number of SNPs deployed (Fig. S7B), which is also strongly suggestive of genome-wide effects contributing toward inbreeding depression rather than the underlying mechanism being a local effect. Finally, an alternative explanation based on population structure seems highly unlikely as we found no evidence for distinct genetic clusters within the RAD dataset (Fig. S8).

Although many reported HFCs yield highly significant P values, the proportion of variance explained tends to be small, typically of the order of 1–5% (34). This could be interpreted as meaning that heterozygosity only weakly affects fitness. Alternatively, heterozygosity could have a large effect on fitness, but because a small panel of microsatellites provides a poor estimate of f , these studies could have limited power to detect such effects. Theory predicts that, if inbreeding is the primary mechanism and thousands of markers could be deployed, the proportion of fitness explained by HFCs would rise considerably (15, 16, 18). This is exactly what we found in wild harbor seals, where the deviance in lungworm infection status explained by heterozygosity increased from around 10% based on 27 microsatellites to nearly 50% with 14,585 SNPs. This is not only qualitatively, but also quantitatively consistent with the theoretical predictions. Under the hypothesis that inbreeding is responsible for HFCs, when switching from microsatellites to RAD SNPs, the variance in infection status explained by heterozygosity ($r^2_{sMLH, infection}$) should increase in proportion to the predicted squared correlation coefficient between heterozygosity and f . Based on the g_2 and variance in sMLH in microsatellites and SNPs estimated in young seals, the expected $r^2_{sMLH, f}$ is 0.148 for microsatellites and 0.709

for SNPs. This expected 4.8-fold difference is in line with the observed 5-fold increase in deviance explained by the HFC. Thus, our results are consistent with inbreeding theory.

Our study clearly demonstrates that at least some HFCs may explain more variation in fitness than previously thought. However, we compared cases with controls, and it is possible that the former could be enriched for a small subset of unusually inbred individuals. Natal philopatry and breeding site fidelity can also be extremely strong in pinnipeds (35, 36) and might combine with polygyny to increase the risk of inbreeding. Moreover, historical changes in the structure of a population, including bottlenecks or population admixture, may also create variance in inbreeding in a broad sense (16). It therefore remains to be seen whether inbreeding depression could be responsible for HFCs more generally, as suggested by Szulkin et al. (16). This represents a fertile area for future research.

Finally, within a genome one expects that heterozygosities should be more correlated between closely linked loci than between unlinked ones (37, 38). Linkage is expected to result in stronger HFCs in some linkage groups than in others (i.e., local effects). However, in our study, models ignoring linkage strongly predict survival and parasite infection status in harbor seals, suggesting that local effects contribute very little to these HFCs, as predicted by theoretical arguments (16). The observed increase in variance explained when thousands of loci are deployed at least rules out the possibility that these relationships arise from local linkage disequilibria between microsatellites and one or a few phenotypically important loci in the absence of genome-wide inbreeding.

In conclusion, the fact that microsatellite heterozygosity is often poorly correlated with f in natural populations has been used as an argument for obtaining more and better pedigrees (39). However, we show that RAD sequencing is capable of generating enough SNP data to accurately estimate inbreeding, in this particular case even more accurately than from a pedigree of reasonable depth. Our approach is also powerful for studying heterozygosity because it can be applied to virtually any species without the need for prior genomic information.

Methods

Oldfield Mouse. A population of oldfield mice was founded at Brookfield Zoo from 26 wild-caught individuals. These mice were paired to produce offspring with a range of inbreeding coefficients (0–0.453) over six generations of laboratory breeding and the resulting pedigree was recorded. Using 179 of these mice, Dasmahapatra et al. (20) reported a significantly negative correlation between pedigree-based inbreeding coefficient and heterozygosity estimated from 12 microsatellite loci. We selected 40 of these individuals at random for RAD sequencing, 36 of which yielded data of sufficient quality for analysis (details in *SI Methods*). The breeding study and the animal care protocols for the oldfield mice were approved by the Institutional Animal Care and Use Committee of the Chicago Zoological Society.

Harbor Seal. Rijks et al. (21) reported a negative association between heterozygosity at 27 microsatellites and lungworm burden in dead harbor seals stranded on the Dutch Wadden Sea coast. However, this was only statistically significant for young seals, defined as being less than 1 y of age, which were more likely to carry lungworms than older seals (21). We therefore RAD sequenced all 43 of the young seals from this study, together with 37 older seals selected at random from the 161 available. The sampling locations of these individuals are shown in Fig. S9. RAD sequence data of adequate quality were obtained for 30 of the young seals, of which 23 were infected with lungworms, plus 30 older seals, none of which carried lungworm infections (*SI Methods*).

RAD Genotyping. RAD libraries were constructed from whole genomic DNA using the protocol of Baird et al. (19) with minor modifications (details in *SI Methods*), and each library was 100-bp paired-end sequenced on an Illumina HiSeq2000 flow cell. Sequences have been deposited in the Short Read Archive (accession no. PRJEB5164). Separately for the mice and seals, Stacks (40) was used to quality filter reads, demultiplex samples, and assemble RAD stacks together with their associated paired-end contigs. After further stringent filtering steps, RAD contigs were used as a reference against which to map

back the raw reads using BWA (41), followed by maximum-likelihood genotype calling using the GATK UnifiedGenotyper (42). Individual heterozygosity and pairwise relatedness were calculated using tag sequences containing only one biallelic SNP per tag. A detailed description of the bioinformatics pipeline used is provided in *SI Methods*.

Data Analyses. For all analyses of the microsatellite and RAD data, individual heterozygosity was expressed as sMLH, which is defined as the total number of heterozygous loci in an individual divided by the sum of average observed heterozygosities in the population over the subset of loci successfully typed in the focal individual (10). Pairwise relatedness (RAD allele sharing) was calculated as the total number of identical alleles between individuals (zero, one, or two per tag) divided by twice the number of tags considered. The two-locus heterozygosity disequilibrium g_2 was measured following David et al. (22) with the method of computation modified to analyze many thousands of loci in a reasonable computing time (details in *SI Methods*). Sensitivity of this estimate to the number of loci was explored by randomly selecting different-sized subsets of loci (between 50 and 15,000) and recalculating g_2 100 times. Distributions of RAD sMLH under the hypothesis $g_2 = 0$ were obtained by shuffling genotypes at each locus randomly

across individuals 1,000 times; as the resulting distribution is very close to normal and the mean is one according to the definition of sMLH, we extracted the variance from these 1,000 simulations and used it to represent the expected distribution as a normal curve. The mean correlation in heterozygosity across loci based on 1,000 random samples was calculated using Rhh (43).

Statistical analyses were conducted using GLMs within R (44). A linear error structure was used for oldfield mouse body mass and a binomial error structure was used for harbor seal age (coded as 0 = young and 1 = old) and lungworm burden (coded as 0 = uninfected and 1 = infected). Where models included multiple predictor variables, standard deletion testing procedures using F tests were used to sequentially remove each term unless doing so significantly reduced the amount of deviance explained (deviance is analogous to sums of squares in standard regression analysis).

ACKNOWLEDGMENTS. This research was supported by a Marie Curie FP7-Reintegration grant within the Seventh European Community Framework Programme (PCIG-GA-2011-303618), a Deutsche Forschungsgemeinschaft standard grant, and a Biotechnology and Biological Sciences Research Council grant (BB/G006903/1).

- Charlesworth D, Charlesworth B (1987) Inbreeding depression and its evolutionary consequences. *Annu Rev Ecol Syst* 18:237–268.
- Jimenez JA, Hughes KA, Alaks G, Graham L, Lacy RC (1994) An experimental study of inbreeding depression in a natural habitat. *Science* 266:271–273.
- Meagher S, Penn DJ, Potts WK (2000) Male-male competition magnifies inbreeding depression in wild house mice. *Proc Natl Acad Sci USA* 97:3324–3329.
- Keller LF (1998) Inbreeding and its fitness effects in an insular population of song sparrows (*Melospiza melodia*). *Evolution* 52:240–250.
- Walling CA, et al. (2013) Inbreeding depression in red deer calves. *BMC Evol Biol* 11:318.
- Hansson B, Westerberg L (2007) Heterozygosity-fitness correlations within inbreeding classes: Local or genome-wide effects? *Conserv Genet* 9(1):73–83.
- David P (1998) Heterozygosity-fitness correlations: New perspectives on old problems. *Heredity* 80:531–537.
- Hansson B, Westerberg L (2002) On the correlation between heterozygosity and fitness in natural populations. *Mol Ecol* 11:2467–2474.
- Coltman DW, Bowen WD, Wright JM (1998) Birth weight and neonatal survival of harbour seal pups are positively correlated with genetic variation measured by microsatellites. *Proc R Soc Lond B Biol Sci* 265:803–809.
- Coltman DW, Pilkington JG, Smith JA, Pemberton JM (1999) Parasite-mediated selection against inbred Soay sheep in a free-living, island population. *Evolution* 53:1259–1267.
- Slate J, Kruuk LEB, Marshall TC, Pemberton JM, Clutton-Brock TH (2000) Inbreeding depression influences lifetime breeding success in a wild population of red deer (*Cervus elaphus*). *Proc R Soc Lond B Biol Sci* 267:1657–1662.
- Seddon N, Amos W, Mulder RA, Tobias JA (2004) Male heterozygosity predicts territory size, song structure and reproductive success in a cooperatively breeding bird. *Proc R Soc Lond B Biol Sci* 271:1823–1829.
- Marshall RC, Buchanan KL, Catchpole CK (2003) Sexual selection and individual genetic diversity in a songbird. *Proc R Soc Lond B Biol Sci* 270:248–250.
- Hoffman JI, Forcada J, Trathan PN, Amos W (2007) Female fur seals show active choice for males that are heterozygous and unrelated. *Nature* 445:912–914.
- Balloux F, Amos W, Coulson T (2004) Does heterozygosity estimate inbreeding in real populations? *Mol Ecol* 13:3021–3031.
- Szulkin M, Bierne N, David P (2010) Heterozygosity-fitness correlations: A time for reappraisal. *Evolution* 64:1202–1217.
- Charlesworth D (2006) Balancing selection and its effects on sequences in nearby gene regions. *PLoS Genet* 2:e64.
- Hoffman JI, Forcada J, Amos W (2010) Exploring the mechanisms underlying a heterozygosity-fitness correlation for canine size in the Antarctic fur seal *Arctocephalus gazella*. *J Hered* 101:539–552.
- Baird NA, et al. (2008) Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PLoS ONE* 3:e3376.
- Dasmahapatra KK, Lacy RC, Amos W (2007) Estimating levels of inbreeding using AFLP markers. *Heredity* 100:286–295.
- Rijks JM, Hoffman JI, Kuiken T, Osterhaus ADME, Amos W (2008) Heterozygosity and lungworm burden in harbour seals (*Phoca vitulina*). *Heredity* 100:587–593.
- David P, Pujol B, Viard F, Castella V, Goudet J (2007) Reliable selfing rate estimates from imperfect population genetic data. *Mol Ecol* 16:2474–2487.
- Miller JM, et al. (2014) Estimating genome-wide heterozygosity: Effects of demographic history and marker type. *Heredity* 112:240–247.
- DeWoody YD, DeWoody JA (2005) On the estimation of genome-wide heterozygosity using molecular markers. *J Hered* 96(2):85–88.
- Santure AW, et al. (2010) On the use of large marker panels to estimate inbreeding and relatedness: Empirical and simulation studies of a pedigreed zebra finch population typed at 771 SNPs. *Mol Ecol* 19:1439–1451.
- Forstmeier W, Schielzeth H, Mueller JC, Ellegren H, Kempenaers B (2012) Heterozygosity-fitness correlations in zebra finches: Microsatellite markers can be better than their reputation. *Mol Ecol* 21:3237–3249.
- Lachance L, Tishkoff SA (2013) SNP ascertainment bias in population genetic analyses: Why it is important, and how to correct it. *Bioassays* 35:780–786.
- Clark AG, Hubisz MJ, Bustamante CD, Williamson SH, Nielsen R (2013) Ascertainment bias in studies of human genome-wide polymorphism. *Genome Res* 15:1496–1502.
- Lynn WJ (2000) Social organization and burrow-site selection of the Alabama beach mouse (*Peromyscus polionotus ammobates*). MSc thesis (Auburn University, Auburn, AL).
- Curtis C, Stewart BS, Karl SA (2009) Pleistocene population expansions of Antarctic seals. *Mol Ecol* 18:2112–2121.
- Matthee CA, Fourie F, Oosthuizen WH, Meyer MA, Tolley KA (2006) Mitochondrial DNA sequence data of the Cape fur seal (*Arctocephalus pusillus pusillus*) suggest that population numbers maybe affected by climatic shifts. *Mar Biol* 148:899–905.
- Kappe AL, Bijlsma R, Osterhaus ADME, VanDelden W, VandeZande L (1997) Structure and amount of genetic variation at minisatellite loci within the subspecies complex of *Phoca vitulina* (the harbour seal). *Heredity* 78:457–463.
- Slate J, Pemberton JM (2002) Comparing molecular measures for detecting inbreeding depression. *J Evol Biol* 15:20–31.
- Chapman JR, Nakagawa S, Coltman DW, Slate J, Sheldon BC (2009) A quantitative review of heterozygosity-fitness correlations in animal populations. *Mol Ecol* 18:2746–2765.
- Hoffman JI, Forcada J (2012) Extreme natal philopatry in female Antarctic fur seals (*Arctocephalus gazella*). *Mamm Biol* 77(1):71–73.
- Hoffman JI, Trathan PN, Amos W (2006) Genetic tracking reveals extreme site fidelity in territorial male Antarctic fur seals *Arctocephalus gazella*. *Mol Ecol* 15:3841–3847.
- Leutenegger AL, et al. (2003) Estimation of the inbreeding coefficient through use of genomic data. *Am J Hum Genet* 73:516–523.
- Hill WG, Weir BS (2011) Variation in actual relationship as a consequence of Mendelian sampling and linkage. *Genet Res* 93(1):47–64.
- Pemberton J (2004) Measuring inbreeding depression in the wild: The old ways are the best. *Trends Ecol Evol* 19:613–615.
- Catchen JM, Amores A, Hohenlohe P, Cresko W, Postlethwait JH (2011) *Stacks*: Building and genotyping loci *de novo* from short-read sequences. *G3* 1(3):171–182.
- Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25:1754–1760.
- McKenna A, et al. (2010) The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res* 20:1297–1303.
- Alho J, Valimäki K, Merilä J (2010) Rhh: An R extension for estimating multilocus heterozygosity and heterozygosity-heterozygosity correlation. *Mol Ecol Res* 10:720–722.
- Ihaka R, Gentleman R (1996) R: A language for data analysis and graphics. *J Comput Graph Stat* 5:299–314.

Supporting Information

Hoffman et al. 10.1073/pnas.1318945111

SI Methods

Restricted Site Associated DNA Library Construction. Total genomic DNA was extracted from mouse skin samples and seal kidney samples stored in 95% (vol/vol) ethanol at -20°C , using a modified phenol-chloroform protocol (1). Eight hundred nanograms of DNA from each sample was individually digested with 20 units SbfI, followed by the ligation of P1 adapters with unique 5-base barcodes for each individual in a restriction site associated DNA (RAD) library. To minimize errors during sequence demultiplexing, at least 2 bases differed between each of the P1 adapter barcodes. Uniquely barcoded samples were pooled and then sheared to ~ 400 bp on a Covaris S2 sonicator. For each library, fragments in the size range 300–700 bp were excised from an agarose gel. Following end repair and A-tailing, a P2 paired-end adapter (P2 top oligo 5'-5Phos/CTCAGGCATCACTCGATTCCCTCCGA-GAACAA-3' and P2 bottom oligo 5'-CAAGCAGAAGACGGCA-TACGACGGAGGAATCGAGTGTGCCTGAG*T-3', where * denotes a phosphorothioate bond; both oligos of the P1 adapters were also modified with a phosphorothioate bond at the same position) was ligated to the size-selected DNA. This template was subjected to 16–17 cycles of PCR enrichment, followed by agarose gel size selection of the 300 to 700-bp fraction. Three oldfield mouse (*Peromyscus polionotus*) and four harbor seal (*Phoca vitulina*) RAD libraries were prepared, each comprising a pool of 14 and 20 individuals, respectively. Each library was paired-end sequenced on an Illumina HiSeq2000 flow cell. DNA quantification was carried out using a Qubit fluorometer (Invitrogen). Agencourt AMPure XP magnetic beads (Beckman Coulter) were used for all reaction clean-up stages.

Bioinformatic Analyses. FastQC (www.bioinformatics.babraham.ac.uk/projects/fastqc/) was used for initial sequence quality assessment. Fluidics problems were encountered during the sequencing of the harbor seal RAD libraries. Consequently, the last 36 bases of the paired-end reads required trimming due to very low sequence quality. These libraries were sequenced a second time and the two datasets combined for subsequent analyses. Stacks process_radtags.pl (2) was used to filter the raw fastq sequences and to demultiplex the samples according to the P1 barcode. At this stage, 6 oldfield mouse and 20 harbor seal samples with very low numbers of sequences were removed from the datasets. The failure of these libraries was due to the inadvertent use of faulty P1 adapters for these samples.

Our pipeline for obtaining single nucleotide polymorphism (SNP) genotypes from the Illumina sequence data involved: (i) clustering of sequences into RAD contigs using Stacks version 0.9999 (2); (ii) using the resulting contigs as a reference genome for mapping the sequences within BWA version 0.6.2 (3); and (iii) SNP calling using the GATK UnifiedGenotyper version 2.1.13 (4). The GATK UnifiedGenotyper (4) uses a Bayesian genotype likelihood model outputting accurate posterior probabilities of there being a segregating variant allele at each locus as well as for the genotype of each sample. Thus, our pipeline allowed us to take advantage of the more sophisticated and statistically more rigorous genotyping and SNP calling framework implemented within the GATK UnifiedGenotyper compared with Stacks.

i) Clustering of sequences into RAD contigs using Stacks. Sequences from all individuals were combined to create a “superparent”. Stacks denovo_map.pl (2) was then used with the superparent acting as a pseudoparent to de novo assemble read 1 sequences into RAD tags. To remove potentially spurious and uninformative tags, the tags generated above were filtered to include only those present

in at least two individuals. The paired-end reads corresponding to each of the remaining tags were assembled into contigs using the Stacks sort_read_pairs.pl and exec_velvet.pl scripts (2). Stacks constructs individual tags based on read 1 sequence similarity but does not take into account the paired-end sequence. Therefore, as an additional quality-control step aimed at eliminating any tags potentially comprising more than one locus, only tags for which a single contig was assembled from the paired-end reads were retained. A reference genome was then constructed from these tag sequences together with their corresponding paired-end contigs, padded out with Ns corresponding to the average size of the RAD library sequenced.

ii) Mapping sequences using the reference genome. The original demultiplexed paired-end fastq files were mapped back to the reference genome using BWA (3) with default parameters. SAMtools (5) was used for SAM and BAM file manipulation. Picard MarkDuplicates version 1.89 (<http://picard.sourceforge.net>) was used to remove PCR duplicates. Individual BAM files for each sample were merged into a single file.

iii) Genotype calling. Genotypes were called using the GATK UnifiedGenotyper (3) with default parameters (6) except with $-\text{hets}$ 0.01 to reflect the higher levels of polymorphism found in these mammals compared with humans. As linked SNPs are non-independent, only tags containing a single polymorphic SNP were retained for subsequent analyses. This measure also guards against the inclusion of false-positive SNPs assembled from paralogous loci (7).

Calculation of RAD-Based Heterozygosity and Relatedness. An individual's heterozygosity was calculated as the total number of heterozygous tags divided by the number of tags for which the individual was called. Because not all individuals were called for the same loci, we then standardized individual heterozygosity values by the mean average observed heterozygosity in the population of the subset of loci successfully typed in the focal individual [standardized multilocus heterozygosity (sMLH)] (8). Pairwise relatedness (RAD allele sharing) was calculated as the total number of identical alleles between individuals (zero, one, or two per tag) divided by twice the number of tags considered.

Filtering of Genotypes. To maximize the signal-to-noise ratio, we explored a range of filtering thresholds based on genotype quality (GQ), low coverage (LC), and mapping quality (MQ). The oldfield mouse pedigree was used to assess the impact of such filtering by measuring the strength of correlation between (i) pedigree-based inbreeding coefficient f and RAD heterozygosity and (ii) pedigree-based relatedness and RAD allele sharing (Figs. S3 and S4, respectively). Studies often use a GQ threshold >30 . However, applied to our data we observed a strong systematic bias whereby individuals with lower sequence depth of coverage, and therefore fewer called SNPs, tended to be called as highly heterozygous. This probably reflects the fact that a heterozygote genotype can be called with higher confidence at low depth of coverage, but to be confident of a homozygote call requires a larger number of reads. To further explore this bias at low sequence coverage, we randomly subsampled the oldfield mouse RAD sequences across all individuals to mimic 50% and 25% of the actual coverage (Figs. S3 and S4, respectively).

Filtering based on MQ thresholds had a negligible impact on the data. However, the r^2 between pedigree-based f and RAD heterozygosity declined with a combination of increasing GQ

and decreasing LC threshold, and this pattern was exacerbated when average sequencing depth was reduced by subsampling the data (Fig. S3). The correlation between pedigree-based relatedness and RAD allele sharing was also optimal for filtering thresholds of $GQ \geq 1$ and a $LC \geq 2$, as was the number of RAD tags retained (Fig. S4). Consequently, to provide the best balance between SNP quality and the number of tags retained for analysis, we filtered genotypes obtained from GATK using a $GQ \geq 1$, corresponding to the maximum-likelihood genotype, and a $LC \geq 2$ to generate the final oldfield mouse SNP dataset. The same filtering criteria were subsequently used for the harbor seals.

Computation of g_2 with Large Numbers of Loci. Notations. In the following we denote H_{ik} an indicator variable that takes a value of 1 if individual k is heterozygous at locus i and 0 otherwise. When there are missing data, some of these values are unknown; in that case we denote \tilde{H}_{ik} a variable that takes a value of 1 if the individual k is known to be heterozygous at locus i and 0 if its genotype at locus i is either unknown or known to be homozygous. x_{ik} denotes a constant that takes a value of 1 if the datum is missing for individual k at locus i and 0 otherwise. The number of individuals in the sample is N , whereas the number of loci is L . We denote $\tilde{h}_k = \sum_{i=1}^L \tilde{H}_{ik}$ the number of known heterozygous loci in an individual k , $\mu_i = E(H_i)$ the expectation of true heterozygosity at locus i in the population, and m_i and m_{ij} the proportions of individuals with missing data at locus i and at both locus i and locus j , respectively. We also define $\tilde{\mu}_i = (1 - m_i)\mu_i$ the expected proportion of individuals that can be successfully scored and found heterozygous in the sample. The m s are considered as constants characteristic of a sample, whereas the μ s are population characteristics. Hats (^) denote estimates based on data from a sample, rather than true values of population parameters.

Estimation of g_2 . The estimates of g_2 presented by David et al. (9) and implemented in the RMES software are (correcting for typographical errors)

$$\hat{g}_2 = \frac{\sum_{i=1}^L \sum_{j \neq i} \left(\sum_{k=1}^N \tilde{H}_{ik} \tilde{H}_{jk} \right)}{\sum_{i=1}^L \sum_{j \neq i} \frac{1}{(N-1)} \left(\sum_{k_1=1}^N \sum_{k_2 \neq k_1} \tilde{H}_{ik_1} \tilde{H}_{jk_2} \right)} - 1 \quad \text{[S1]}$$

in the absence of missing data, which becomes

$$\hat{g}_2 = \frac{\sum_{i=1}^L \sum_{j \neq i} \frac{1}{N(1 - m_i - m_j + m_{ij})} \left(\sum_{k=1}^N \tilde{H}_{ik} \tilde{H}_{jk} \right)}{\sum_{i=1}^L \sum_{j \neq i} \frac{1}{N(N-1)(1 - m_i - m_j + m_i m_j) - N(m_{ij} - m_i m_j)} \left(\sum_{k_1=1}^N \sum_{k_2 \neq k_1} \tilde{H}_{ik_1} \tilde{H}_{jk_2} \right)} - 1$$

in the presence of missing data.

These estimates are impractical when the number of loci is high because of the double summations over all pairs of loci. With 15,000 loci, the double summations take of the order of 0.2×10^9 computation steps (which then have to be multiplied by N^2 as there are also double summations over individuals). To reduce

computation time, we can look for an estimate of g_2 that takes a more computationally tractable form. The basic assumption behind this computation (which also underlies the previous estimates) is that the distribution of true heterozygosity is the same in missing data as in nonmissing data. In such conditions it can be shown that

$$g_2 = \frac{1 + (B - C)/A}{1 + \bar{\alpha}} - 1,$$

where

$$A = \sum_{i=1}^L \sum_{j \neq i} \tilde{\mu}_i \tilde{\mu}_j = \left(\sum_{i=1}^L \tilde{\mu}_i \right)^2 - \sum_{i=1}^L \tilde{\mu}_i^2,$$

$$B = VAR(\tilde{h}_k)$$

$$C = \sum_{i=1}^L VAR(\tilde{H}_i)$$

$$\bar{\alpha} = \left[\sum_{i=1}^L \sum_{j \neq i} \tilde{\mu}_i \tilde{\mu}_j \alpha_{ij} \right] / A,$$

with

$$\alpha_{ij} = \frac{m_{ij} - m_i m_j}{(1 - m_i)(1 - m_j)}.$$

The α s represent the extent to which missing loci are clustered within individuals; in the absence of clustering (i.e., missing data occur independently at all loci), they would be zero; however, it is possible that some individuals may have more missing data than others on average, for example because they would have a lower coverage in the RAD sequences. $\bar{\alpha}$ is the weighted average of the α_{ij} s.

Unbiased estimators of the $\tilde{\mu}_i$ s are simply found by averaging over individuals as $\hat{\tilde{\mu}}_i = \frac{1}{N} \sum_{k=1}^N \tilde{H}_{ik}$.

Unbiased estimators of A , B , and C can then be obtained as

$$\hat{A} = \frac{N}{N-1} \left[\left(\sum_{i=1}^L \hat{\mu}_i \right)^2 - \sum_{i=1}^L \hat{\mu}_i^2 \right] - \frac{\hat{J}}{N-1},$$

$$\text{with } \hat{J} = \frac{1}{N} \sum_{k=1}^N \hat{h}_k^2 - \sum_{i=1}^L \hat{\mu}_i,$$

$$\hat{B} = \frac{1}{N-1} \left[\sum_{k=1}^N \hat{h}_k^2 - \frac{1}{N} \left(\sum_{k=1}^N \hat{h}_k \right)^2 \right]$$

$$\hat{C} = \frac{N}{N-1} \left(\sum_{i=1}^L \hat{\mu}_i - \sum_{i=1}^L \hat{\mu}_i^2 \right).$$

All these equations do not require double summation over loci and can therefore be computed in a reasonable time. However, computing $\bar{\alpha}$ in principle requires this double summation. To avoid this, one can assume that the α_{ij} s do not vary a lot between pairs of loci. With this approximation we obtain the following estimate of $\bar{\alpha}$,

$$\hat{\bar{\alpha}} = \frac{\sum_{k=1}^N \hat{M}_k - N \left(\sum_{i=1}^L \hat{\mu}_i \frac{m_i}{1-m_i} \right)^2 + N \sum_{i=1}^L \left(\hat{\mu}_i \frac{m_i}{1-m_i} \right)^2}{(N-1)\hat{A} + \hat{J}},$$

in which

$$\hat{M}_k = \left(\sum_{i=1}^L \frac{\hat{\mu}_i x_{ik}}{1-m_i} \right)^2 - \sum_{i=1}^L \left(\frac{\hat{\mu}_i x_{ik}}{1-m_i} \right)^2.$$

This does not require a double summation over loci and the final estimate of g_2 reads

$$\hat{g}_2 = \frac{1 + [\hat{B} - \hat{C}] / \hat{A}}{1 + \hat{\bar{\alpha}}} - 1.$$

This estimate is slightly biased because the ratio of expectations differs from the expectation of a ratio; however, with reasonable conditions (say, when individuals have on average 10 or more successfully scored, heterozygous loci), the bias is very small (9). The SD can be obtained by bootstrapping over individuals. All of the computations were done using a Mathematica program, available upon request (a Windows executable will be made available in the near future).

Bayesian Analysis of Population Structure. Population structure can potentially generate spurious associations between heterozygosity and fitness (10, 11). Consequently, we used the program Structure 2.3.4 (12) to conduct a Bayesian cluster analysis of the harbor seal RAD dataset, comprising 60 individuals genotyped at 14,585 SNPs. Structure uses a maximum-likelihood approach

to determine both the most likely number of distinct genetic groups (K) in a sample and the probability of membership of every individual to each of these K groups. We ran three independent runs for $K = 1-5$ using 100,000 Markov chain Monte Carlo iterations after a burn-in of 100,000 steps and specifying an admixture model with correlated allele frequencies. To be able to detect even a very weak signature of population structure, we implemented the LOCPRIOR model, using age class as prior information to assist the clustering (13). LOCPRIOR tends to outperform the standard model when populations are weakly differentiated, providing more accurate estimates of K together with improved group membership coefficients.

SI Results

Oldfield Mouse. RAD sequencing of 36 oldfield mice generated 265 million paired-end reads, of which 231 million contained appropriate barcodes and the RAD restriction site and passed initial quality filtering, resulting in an average of 6.4 million paired-end reads per individual, varying between 2.5 million and 11 million (Fig. S1). These reads were assembled into 79,360 contigs, which, after eliminating those with multiple paired-end contigs, were reduced to 63,129. Average contig length including the paired end was 479 bp (± 117 -bp SD). A total of 16,060 contigs (25.4%) mapped to the mouse genome, using an e -value threshold of $1e^{-10}$ to reveal a broad genomic distribution (Fig. S2A). Following SNP calling and filtering as described in *SI Methods*, 13,198 RAD tags were retained for downstream analysis, each containing a single biallelic SNP. Most of these SNPs were called in the majority of individuals (Fig. S5).

Harbor Seal. RAD sequencing of 60 harbor seals generated 374 million paired-end reads, of which 280 million contained appropriate barcodes and the RAD restriction site and passed initial quality filtering, resulting in an average of 3.6 million paired-end reads per individual, varying between 0.6 and 9.6 million (Fig. S1). These data were assembled into 126,121 contigs, of which 83,148 were retained after filtering out those with multiple paired-end contigs. Average contig length including the paired end was 538 bp (± 135 -bp SD). A broad genomic distribution was inferred by mapping 44,961 contigs (35.6%) to the dog genome (Fig. S2B). Following SNP calling and filtering, 14,585 RAD tags were retained for downstream analyses, most of which were present in the majority of sequenced individuals (Fig. S5).

Differences in heterozygosity between old and young seals could potentially arise if the two age classes originated from separate populations rather than from the same panmictic population. Consequently, we subjected our harbor seal RAD dataset to Bayesian cluster analysis, using the program Structure (12). The average log-likelihood value increased from $K = 1$ to $K = 2$ and thereafter leveled off (Fig. S8A). Although this appears to indicate support for the presence of more than one genetic cluster, inspection of the membership probabilities to each of the inferred clusters shows that no signal of population structure is present, with all individuals being predominantly assigned to a single cluster (each bar is almost entirely blue in color) and the remaining clusters making negligible contributions (Fig. S8 B-E).

1. Sambrook J, Fritsch EF, Maniatis T (1989) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab Press, Cold Spring Harbor, NY) 2nd Ed.
2. Catchen JM, Amores A, Hohenlohe P, Cresko W, Postlethwait JH (2011) *Stacks: Building and genotyping loci de novo from short-read sequences*. *G3* 1(3):171-182.
3. Li H, Durbin R (2010) Fast and accurate long-read alignment with Burrows-Wheeler Transform. *Bioinformatics* 26:589-595.
4. McKenna A, et al. (2010) The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res* 20:1297-1303.
5. Li H, et al. (2009) The Sequence alignment/map (SAM) format and SAMtools. *Bioinformatics* 25:2078-2079.
6. DePristo M, et al. (2011) A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat Genet* 43:491-498.

7. Sanchez CC, et al. (2009) Single nucleotide polymorphism discovery in rainbow trout by deep sequencing of a reduced representation library. *BMC Genomics* 10:599.
8. Coltmann DW, Pilkington JG, Smith JA, Pemberton JM (1999) Parasite-mediated selection against inbred Soay sheep in a free-living, island population. *Evolution* 53(4):1259-1267.
9. David P, Pujol B, Viard F, Castella V, Goudet J (2007) Reliable selfing rate estimates from imperfect population genetic data. *Mol Ecol* 16:2474-2487.
10. Slate J, Pemberton JM (2006) Does reduced heterozygosity depress sperm quality in wild rabbits (*Oryctolagus cuniculus*)? *Curr Biol* 16:R790-R792.
11. Luquet E, et al. (2011) Heterozygosity-fitness correlations among wild populations of European tree frog (*Hyla arborea*) detect fixation load. *Mol Ecol* 20:1877-1887.

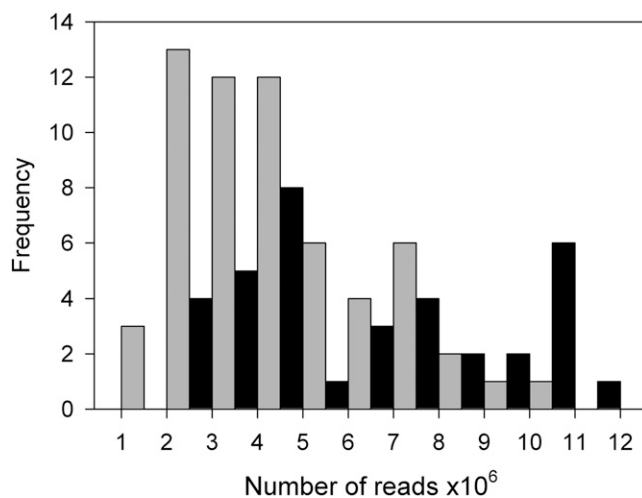


Fig. S1. Distribution of the number of RAD sequence reads obtained across samples. Oldfield mice and harbor seals are denoted by solid and shaded bars, respectively.

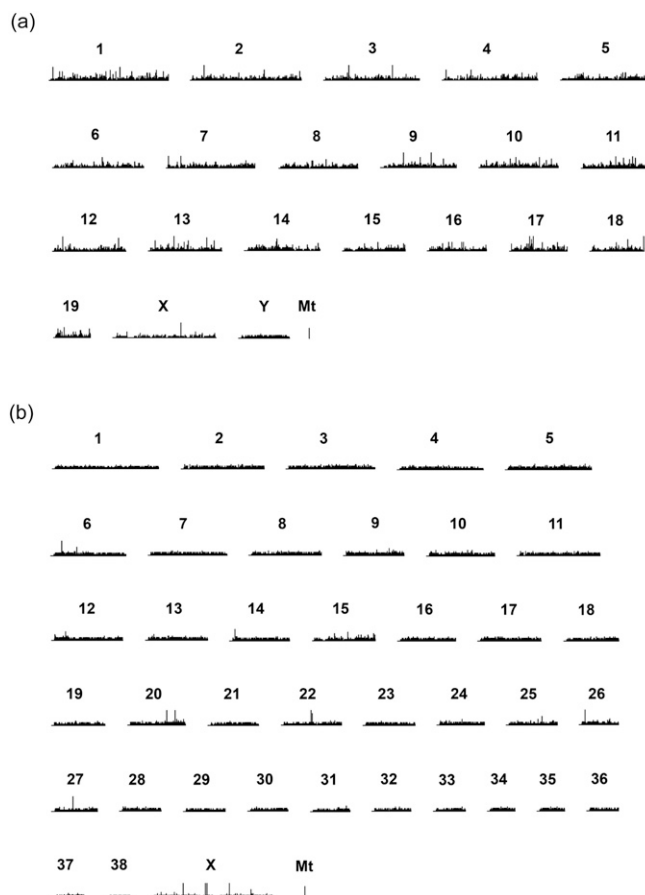


Fig. S2. Inferred chromosomal distributions of (A) oldfield mouse and (B) harbor seal contigs (details in *SI Methods*). The oldfield mouse contigs were BLASTed against the mouse (*Mus musculus*) genome and the harbor seal contigs were BLASTed against the dog (*Canis familiaris*) genome, using an e-value cutoff of $1e^{-10}$. Chromosomal distributions are based on a bin size of 1,000 bp with the x axis being scaled relative to the largest chromosome and the maximal y axis being 10 contigs per bin.

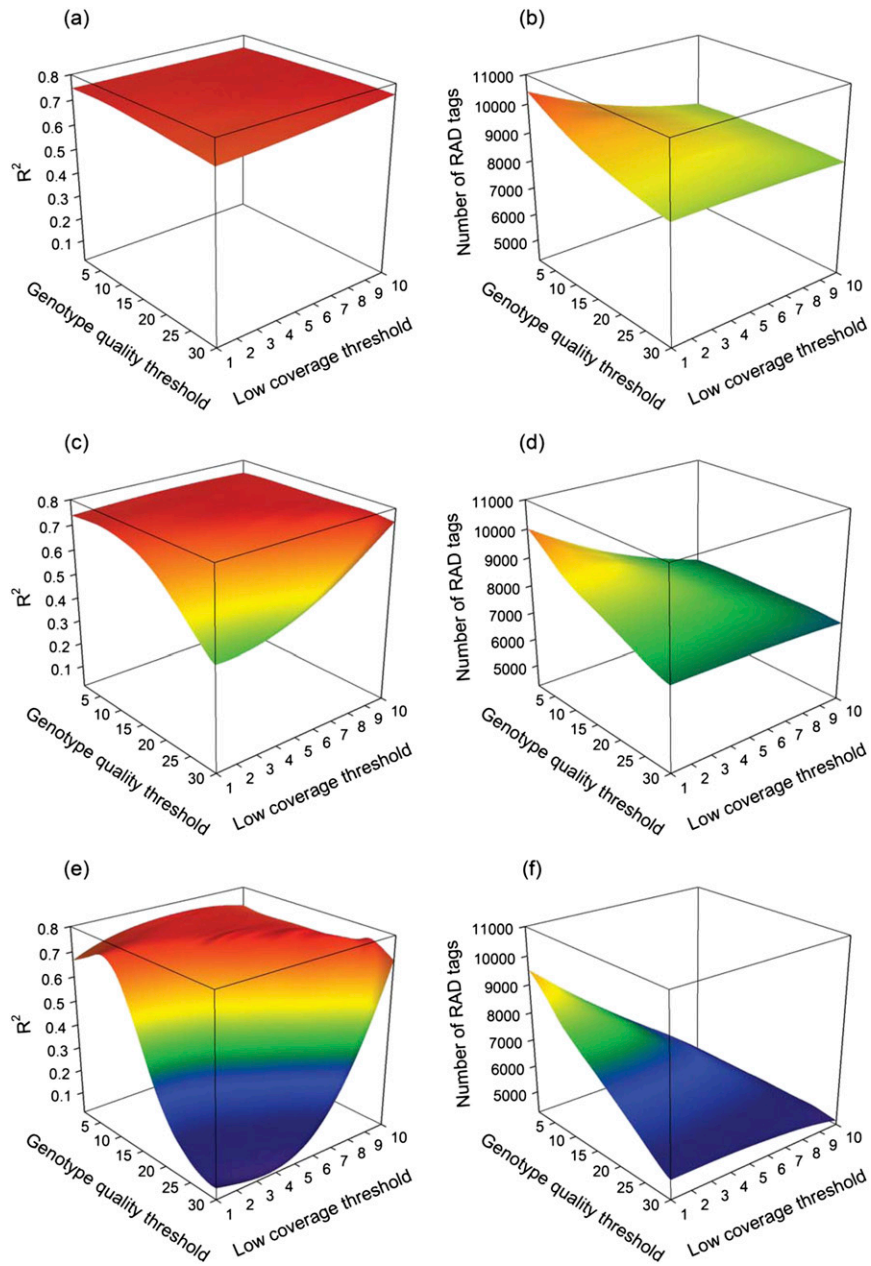


Fig. S3. The effect of varying two genotype filtering thresholds, genotype quality and low-coverage cutoff, on the strength of the relationship between pedigree-based inbreeding coefficient f and RAD heterozygosity in the oldfield mouse. *A*, *C*, and *E* show the r^2 between pedigree f and RAD heterozygosity based on 100%, 50%, and 25% of the sequence data, respectively, across all 36 individuals. *B*, *D*, and *F* show the corresponding median numbers of RAD tags retained for analysis. Local regression, implemented using the “locfit” package in R, was used to fit smoothed splines to the raw datasets.

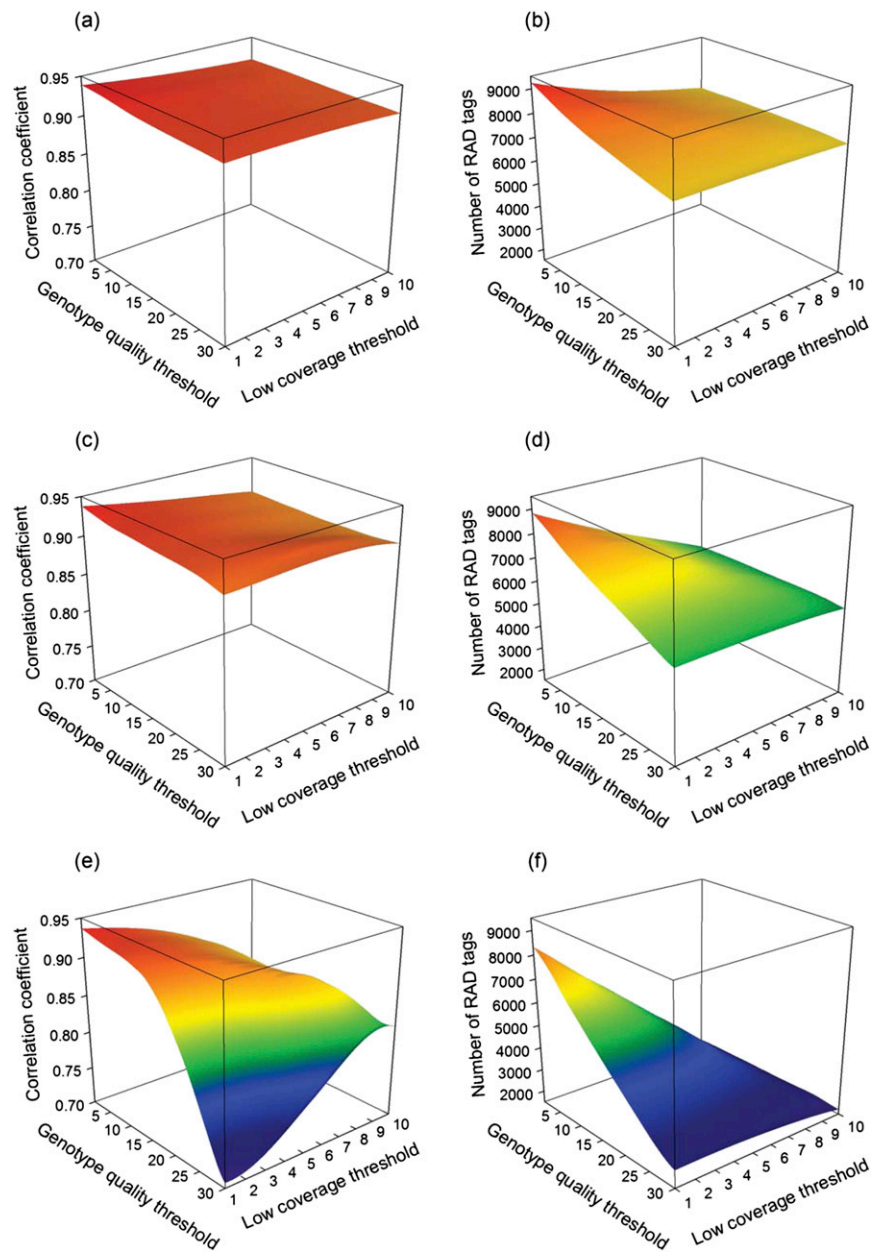


Fig. S4. The effect of varying two genotype filtering thresholds, genotype quality and low-coverage cutoff, on the strength of the relationship between pedigree-based relatedness and RAD allele sharing in the oldfield mouse. *A*, *C*, and *E* show the correlation coefficient between the two relatedness measures based on 100%, 50%, and 25% of the sequence data, respectively. *B*, *D*, and *F* show the corresponding median numbers of RAD tags retained for analysis. Note that fewer RAD tags in general are retained for allele sharing because each tag needs to be called in more than one individual to be counted. Local regression, implemented using the *locfit* package in R, was used to fit smoothed splines to the raw datasets.

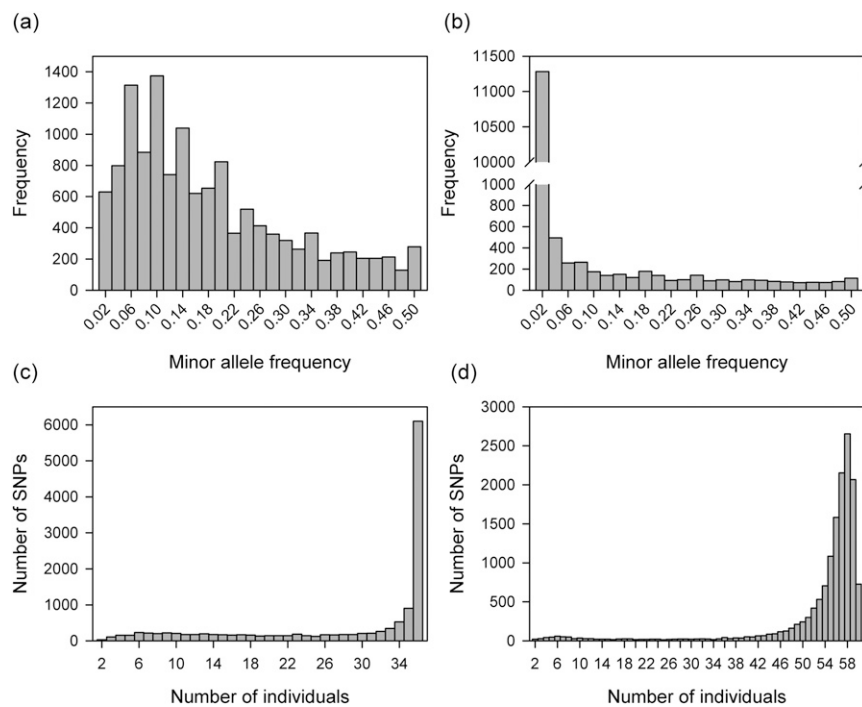


Fig. S5. Summary of 13,198 SNPs called in the oldfield mouse (*A* and *C*) and 14,585 SNPs called in the harbor seal (*B* and *D*). *A* and *B* show the distribution of SNP coverage across individuals. The majority of SNPs were called in most of the individuals. *C* and *D* show corresponding minor allele frequency (MAF) distributions.

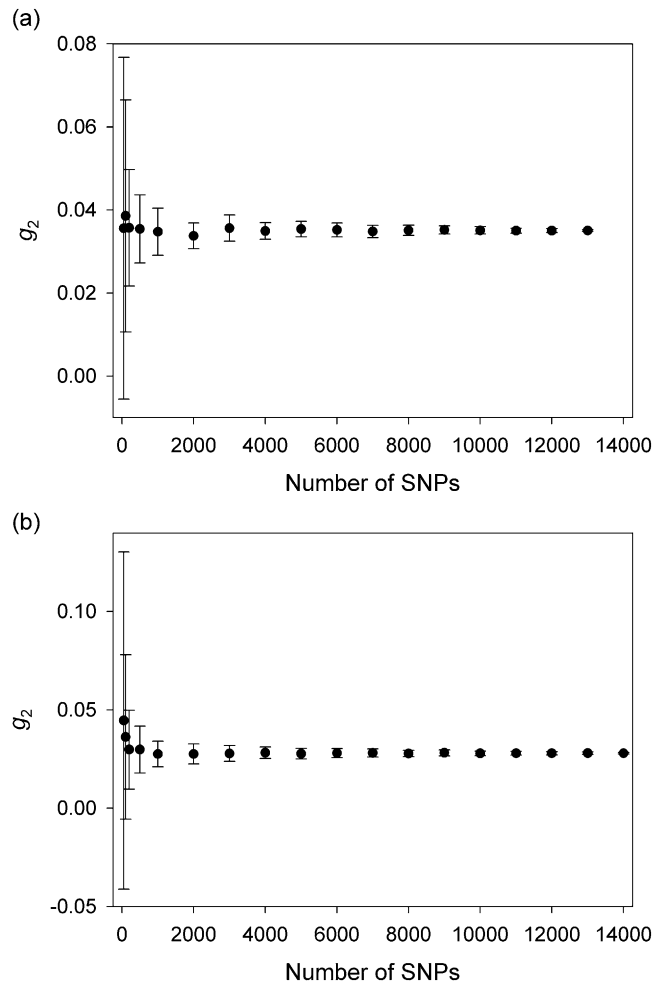


Fig. S6. Relationship between the number of randomly subsampled SNPs and g_2 (\pm SD) in (A) oldfield mice and (B) harbor seals.

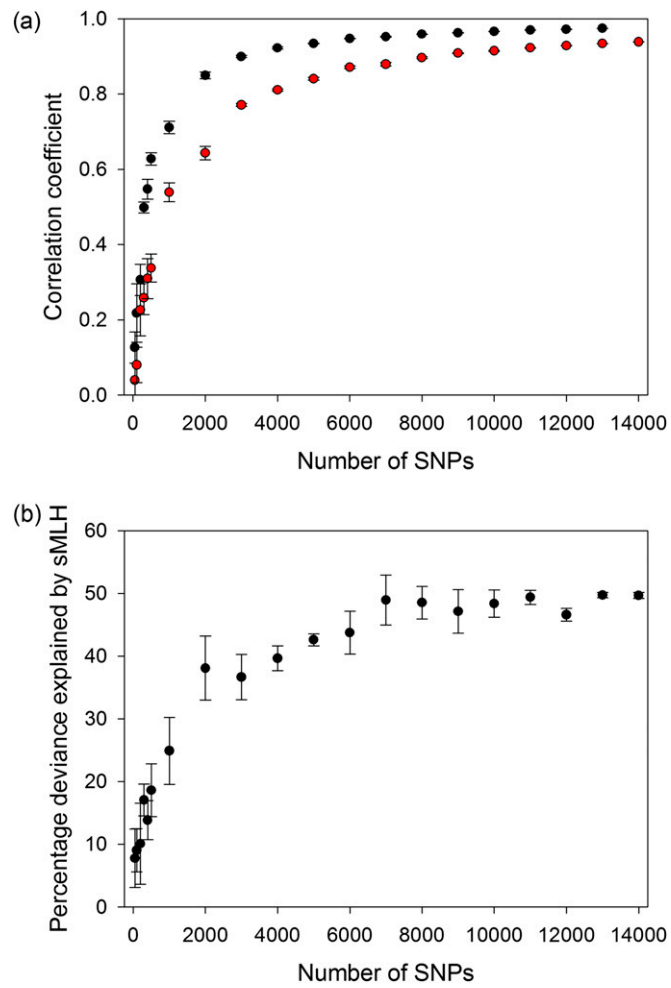


Fig. 57. Relationship between the number of randomly subsampled SNPs and (A) the mean (\pm SEM) correlation in sMLH among markers, with black circles denoting oldfield mice and red circles harbor seals, and (B) the mean (\pm SEM) percentage deviance explained by sMLH in the generalized linear model (GLM) of parasite infection in harbor seals.

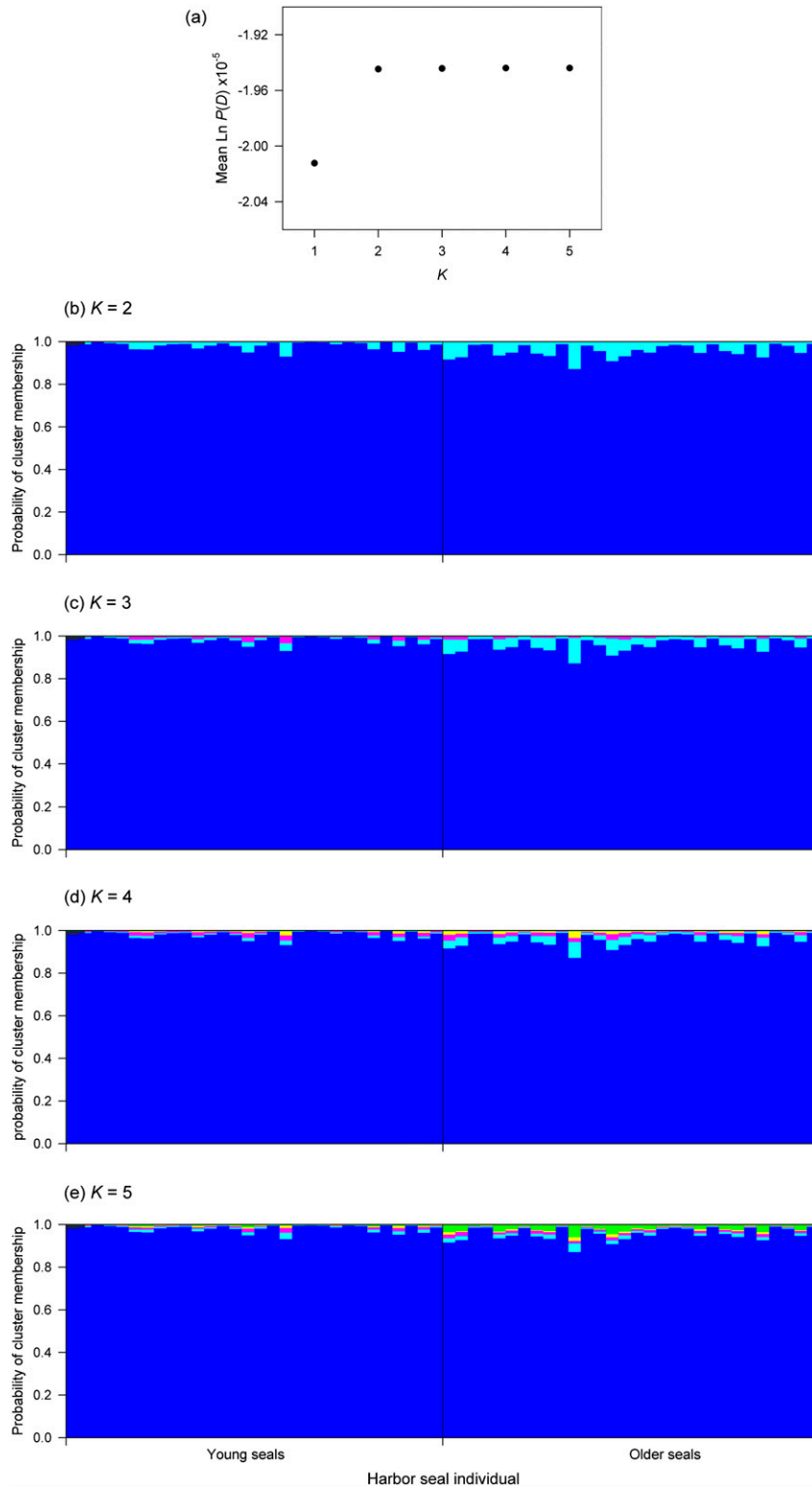


Fig. 58. Results of the Structure analysis of the harbor seal RAD dataset. (A) Average log-likelihood value values based on three replicates for each value of K , the hypothesized number of clusters in the data. (B–E) Clustering results shown separately for $K = 2$ –5. Each individual is represented by a vertical bar partitioned into different segments, the lengths of which indicate the probability of membership in the different clusters.

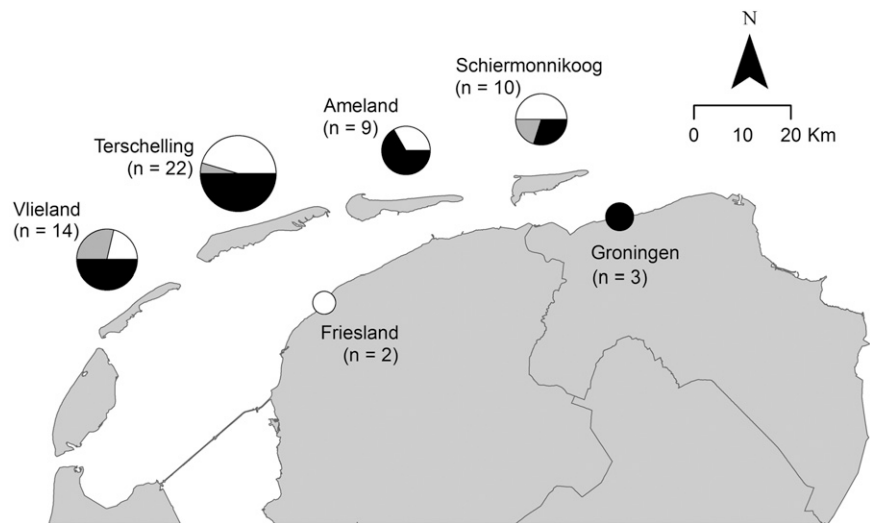


Fig. S9. Spatial distribution of the harbor seals in our sample set. Young seals with lungworm, young seals without lungworm, and old seals without lungworm are denoted by open, shaded, and solid sections, respectively. The diameter of each pie chart corresponds to the number of seals tested per location.