Quantification of Inbreeding Due to Distant Ancestors and Its Detection Using Dense Single Nucleotide Polymorphism Data

Matthew C. Keller,[*,†,1] **Peter M. Visscher,**[‡] **and Michael E. Goddard**[§,**]

*Department of Psychology and Neuroscience and †Institute for Behavioral Genetics, University of Colorado, Boulder, Colorado, 80309, ‡Queensland Statistical Genetics Laboratory, Queensland Institute for Medical Research, Brisbane, 4029, Queensland, Australia, §Faculty of Land and Food Resources, University of Melbourne, Melbourne, 3001, Victoria, Australia, and **Department of Primary Industries, Melbourne, 3001, Victoria, Australia

ABSTRACT Inbreeding depression, which refers to reduced fitness among offspring of related parents, has traditionally been studied using pedigrees. In practice, pedigree information is difficult to obtain, potentially unreliable, and rarely assessed for inbreeding arising from common ancestors who lived more than a few generations ago. Recently, there has been excitement about using SNP data to estimate inbreeding (F) arising from distant common ancestors in apparently "outbred" populations. Statistical power to detect inbreeding depression using SNP data depends on the actual variation in inbreeding in a population, the accuracy of detecting that with marker data, the effect size, and the sample size. No one has yet investigated what variation in F is expected in SNP data as a function of population size, and it is unclear which estimate of F is optimal for detecting inbreeding depression. In the present study, we use theory, simulated genetic data, and real genetic data to find the optimal estimate of F, to quantify the likely variation in F in populations of various sizes, and to estimate the power to detect inbreeding depression. We find that F estimated from runs of homozygosity (F_{roh}), which reflects shared ancestry of genetic haplotypes, retains variation in even large populations (*e.g.*, SD = 0.5% when N_e = 10,000) and is likely to be the most powerful method of detecting inbreeding effects from among several alternative estimates of F. However, large samples (*e.g.*, 12,000–65,000) will be required to detect inbreeding depression for likely effect sizes, and so studies using F_{roh} to date have probably been underpowered.

INBREEDING occurs when mates are more closely related than expected if chosen at random in the population. Most cultures have taboos against close inbreeding (Brown 1991) and most nonhuman animals appear to avoid it, ostensibly as an adaptation to prevent its deleterious effects (Pusey and Wolf 1996). Inbred offspring tend to have higher rates of congenital disorders and lower survival rates and fertility. This phenomenon—called inbreeding depression—has since been confirmed by hundreds of scientific investigations on plants and animals (Roff 1997). The magnitude of the effect appears to be related to the strength of directional selection on the trait. Fitness traits such as survival, reproduction,

resistance to disease, and predator avoidance are more affected by inbreeding than are traits likely to be under less intense, directional selection (DeRose and Roff 1999). Interestingly, there are numerous reports of inbreeding effects on human complex traits such as heart disease (Shami *et al.* 1991), hypertension (Rudan *et al.* 2003b), osteoporosis (Rudan *et al.* 2004), cancer (Lebel and Gallagher 1989), IQ (Morton 1979; Afzal 1988), and psychiatric disorders (Abaskuliev and Skoblo 1975; Gindilis *et al.* 1989; Bulayeva *et al.* 2005), suggesting that these traits or the genetic variants underlying them have been subject to natural selection ancestrally.

Two major hypotheses have been forwarded to explain why inbreeding depression occurs (Wright 1977). The partial dominance hypothesis focuses on the role of homozygosity of rare, recessive/partially recessive deleterious mutations. Deleterious mutations constantly arise in the population and selection quickly purges the most additive and dominant ones, leaving the segregating pool of deleterious mutations enriched with partially recessive ones because selection against

recessive mutations is inefficient. When such mutations meet in homozygous form, such as following inbreeding, their full deleterious effects are exposed. Alternatively, the overdominance hypothesis posits that inbreeding depression is caused by a reduction in heterozygosity of common alleles maintained at equilibrium at loci governed by heterozygote advantage. Both mechanisms may play a role in inbreeding depression effects, but the partial dominance hypothesis enjoys the strongest empirical support to date (Charlesworth and Charlesworth 1999; Charlesworth and Willis 2009).

Estimation of Inbreeding

The inbreeding coefficient of an individual, F, is one of the central parameters in population genetics theory. It is defined as the probability that two randomly chosen alleles at a homologous locus within an individual are *identical by descent* (IBD) with respect to a base (reference) population in which all alleles are independent; that is, the alleles are identical because they are passed down from a common ancestor (Wright 1922). Homozygosity caused by two IBD genomic segments is termed autozygosity, as opposed to allozygosity, which is homozygosity produced by alleles that are *identical by state*. F is therefore an estimate of genome-wide autozygosity.

Traditionally, F has been estimated using known pedigrees (F_{ped}), typically using a path coefficient method developed by Wright (1922). In practice, pedigree information is difficult and costly to obtain, potentially unreliable (*e.g.*, due to problems with accurate recording of ancestry), and rarely assessed for inbreeding arising from common ancestors who lived more than three or four generations in the past. Although autozygosity caused by common ancestors living more than three generations ago contributes very little variation to F_{ped}, it can contribute substantially to variation in segments of the genome that are autozygous (Stam 1980). Moreover, F_{ped} is an expectation of the proportion of the genome that is autozygous, but there is much variation around this expectation due to the stochastic nature of recombination. For example, the percentage of the genome autozygous among progeny of first cousins averages 6.25%, but the standard deviation of this is $\pm 2.4\%$ (Franklin 1977; Hill and Weir 2011).

For these reasons, there has recently been excitement about using dense marker data to estimate F arising from even very ancient inbreeding (Leutenegger *et al.* 2003; Carothers *et al.* 2006; Gibson *et al.* 2006; Li *et al.* 2006; Woods *et al.* 2006; McQuillan *et al.* 2008). Such genomic estimates of F potentially have several advantages over F_{ped}. First, whereas F_{ped} is an expectation of genome-wide autozygosity, by directly measuring homozygosity, genomic estimates of F can potentially estimate the actual percentage of the genome that is autozygous more accurately. Second, genomic estimates of F incorporate autozygosity arising from very distant common ancestors (*e.g.*, 50+ generations ago). Third, genomic estimates of F can be estimated in any

sample that has marker data collected on it, including samples for which pedigree information is difficult or impossible to collect. Fourth, whereas all estimates of F are genome-wide estimates of autozygosity, genomic estimates of F can be altered to allow for the possibility of testing whether an effect of F is distributed evenly across the genome or whether the signal comes from specific genomic locations (*e.g.*, by obtaining separate F estimates for different chromosomes). Fifth, in certain species (*e.g.*, humans), individuals who inbreed may not be a representative sample of the population, and thus putative inbreeding effects may be due to "third variable", nongenetic reasons. Such third variable explanations are less likely in populations where inbreeding is likely to be distant and unintended. Finally, given the decreasing price in genome-wide SNP data, it is likely that genomic estimates of F are less expensive to collect than is F_{ped}, which requires either intensive observation in the field or (in humans) extensive interviews to obtain pedigrees from both parents.

A potential drawback to using genomic estimates of F is that their behavior in populations with different levels of inbreeding is not well characterized, and it is therefore unclear which estimates should be preferred under which situations. Furthermore, there may not be enough variation in genomic estimates of F in unselected ("outbred") samples to detect inbreeding effects with statistical significance.

Current Study

The current study has three main goals. First, there are several potential ways to estimate F: from pedigrees, on a marker-by-marker basis, and from runs of homozygosity. Using simulated data sets that have realistic patterns of molecular variation and linkage disequilibrium, we seek to understand which of these estimates of F are optimal for detecting inbreeding depression and whether this answer depends on the level of inbreeding (assessed using different population sizes) in the population. Second, we are interested in comparing the variance, error, and intercorrelations of the various estimates of F arising from populations of different sizes and that followed population expansions similar to those that occurred in recent human history. Third, on the basis of the two results above, we estimate the likelihood of detecting inbreeding depression given the predicted variances in the optimal estimate(s) of F. In this report, we use the human genome and population history to guide our simulation because much is known about these parameters in humans and because there has been much interest in detecting the effects of autozygosity on human traits. Nevertheless, as discussed below, the findings of this report readily extend to nonhuman animal populations as well.

Methods

Simulation of sequence and SNP data

We simulated populations of different sizes using the Fregene forward-time simulation program (Chadeau-Hyam

et al. 2008). Fregene simulates the evolutionary process of genetic sequencing data in a population following the Fisher–Wright model (a monoecious, diploid, randomly mating population that evolves across nonoverlapping generations). Mating was random except that selfing was not allowed. It should be noted that variation of all estimates of F will be different if mating is not random. Nonrandom mating typically increases variation in F *as does a reduction in N_e*. Therefore smaller effective population sizes could be used as a proxy for studying variation of estimates of F in the presence of nonrandom mating.

The effective population size (N_e) of humans has been estimated at ~10,000 on the basis of molecular variation (Takahata *et al.* 1995) and to have passed through a bottleneck, reducing the population sizes to ~3000 for Caucasians and ~8000 for Africans, on the basis of linkage disequilibrium patterns (Tenesa *et al.* 2007). In the present study, we simulated effective population sizes of $N_e = 100$, 1000, and 10,000. Simulating larger population sizes was not computationally feasible. Each individual's genome was composed of two homologous chromosomes of length 100 Mb. We chose 100-Mb–length genomes for reasons of computational feasibility and because 100 Mb is roughly the size of a typical mammalian chromosome. Fregene's recombination model allows for crossover rates that vary along the chromosome at both broad scales (regions of several megabases in length that differ in background recombination rates) and fine scales (corresponding to recombination hotspots of ~2 kb in length). Recombination rates for the present set of simulations averaged $\nu = 1.1 \times 10^{-8}$ per site per generation, 80% of recombination events occurred in hotspot regions, and the average distance between hotspots was 8500 bp. Mutations arose at rate $\mu = 2.3 \times 10^{-8}$ per site per generation, and all mutations were neutral with respect to fitness. New mutations occurred uniformly and at random across the genome; thus mutations could arise at an already polymorphic site, allowing both "double hit" mutations (if occurring at an ancestral allele) and "back" mutations (if occurring at a derived allele), although these types of mutations were extremely rare. These parameters, the defaults in Fregene, were based on estimated per site recombination and mutation rates in human populations (Chadeau-Hyam *et al.* 2008).

To reduce the computational time taken by our simulations, we used Fregene's *scale_exp* option with the scaling parameter $\lambda = 20$. This reduced the population size and number of generations 20-fold but increased ν and μ 20-fold, which kept the population mutation parameter, $\theta = 4N_e\mu$ and the population linkage parameter, $\rho = 4N_e\nu$, constant, thereby mimicking the degree of variation and linkage disequilibrium in nonscaled populations but decreasing the computational time by over an order of magnitude. Populations evolved for 100,000 generations (*i.e.*, across 5000 loops once scaled), ensuring that they reached drift–mutation–recombination equilibrium (Chadeau-Hyam *et al.* 2008).

One drawback with scaling in Fregene is that the mixture of rare to common variants is slightly inaccurate. For example, with $\lambda = 20$ and $N_e = 1000$, no variant can have a minor allele frequency <0.01 (rather than 0.0005, which it would be if unscaled). To minimize this effect, at the end of the 100,000 generations (5000 loops), we ran Fregene with no scaling for an additional 1000 generations (1000 loops) for each population. The additional unscaled generations allowed polymorphism levels (as judged by heterozygosity) to build up to expected levels. The observed levels of heterozygosity in the three populations at the end of the simulation (9.3×10^{-4}, 9.3×10^{-5}, 9×10^{-6}) conformed closely to $4N_e\mu/(4N_e\mu + 1)$, as predicted by theory (9.2×10^{-4}, 9.2×10^{-5}, 9×10^{-6}).

To investigate the effects of population size on various estimates of F, we needed to (a) obtain information on pedigrees going back several generations, (b) derive SNP data from the sequences in the samples, and (c) create samples of equal numbers of individuals, despite the varying population sizes. To accomplish the first of these goals, we modified the Fregene program to track and write out pedigree information for each individual going back five generations (as we show, going back farther than this was unnecessary). To accomplish the second goal of creating SNP data, we selected all variants with minor allele frequency (MAF) > 0.05 from the sequence data that Fregene outputs at the final generation. This resulted in ~2700 SNPs (1 SNP per 36 kb) in samples from the $N_e = 100$ population, ~27,000 SNPs (1 SNP per 3.7 kb) in samples from the $N_e = 1000$ population, and ~274,000 SNPs (1 SNP per 360 bp) in samples from the $N_e = 10,000$ population. These differences in SNP densities reflect the larger numbers of variants that naturally occur in larger populations (Crow and Kimura 1970).

The third goal, to create equal experimental sample sizes for each population size, was motivated by the fact that we did not want our results across different population sizes to be confounded by differing sample sizes. To accomplish the goal of analyzing adequately large samples of equal sizes across the three different population sizes, for the $N_e = 100$ population, 100 subpopulations (each of $N_e = 100$) split off from the original $N_e = 100$ population and evolved independently for 50 generations. Groups of 10 of these subpopulations were randomly placed together (without replacement) to create 10 samples of size $n = 1000$ each. Because all 100 subpopulations evolved from the same progenitor population, the common variants were generally the same between subpopulations, allowing SNP sets to be created from the $n = 1000$ samples. Thus, despite having the same levels of inbreeding as in a population of size $N_e = 100$, each of 10 replicate samples consisted of 1000 individuals. Similarly, for the $N_e = 1000$ population, 10 subpopulations of size $N_e = 1000$ split off from the original population (of size $N_e = 1000$) and evolved independently for 50 generations; each of these populations was converted to 10 samples of size $n = 1000$ each at the end of the 50 generations. Finally, the $N_e = 10,000$ population was split into 10 replication samples of $n = 1000$ individuals each and evolved for 50 generations. In this way, 10 samples of size $n = 1000$ were
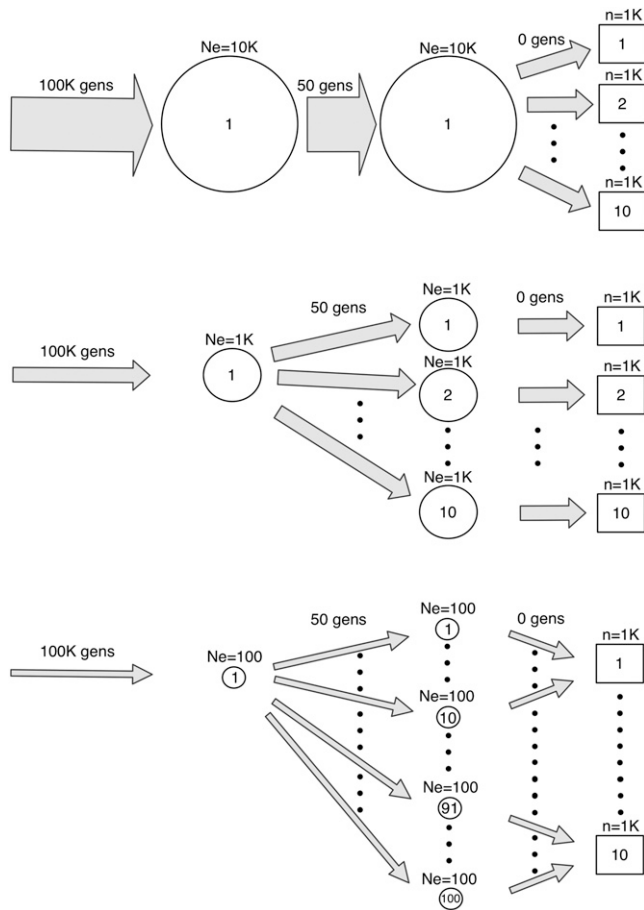
Figure 1 Procedure for deriving 10 samples from each of three effective population sizes. Circles represent populations, arrows represent evolution and splitting/combining of each population, and squares represent samples of size 1000 derived from each population. The sizes of the shapes correspond to population (circles and arrows) or to sample (squares) sizes. See text for details.

derived from the different population sizes (see Figure 1). By evolving independently for 50 generations, each subpopulation (and therefore each sample) had independent breeding (and inbreeding) histories within a 50-generation time frame.

To study the effects of population expansion on estimates of F, we allowed a population of size $N_e = 100$ that had reached drift–mutation equilibrium to expand instantaneously to $N_e = 10,000$ and then to evolve at this size for 400 generations [~11,400 years for humans (Fenner 2005)]. Computational limitations disallowed investigating larger and more realistic expansions (*e.g.*, $N_e = 10,000$–1,000,000), but conclusions from studying smaller-scale expansions should apply to larger expansions as well. We wrote out pedigree, sequencing, and SNP data for 10 samples of size $n = 1000$, as described above, at generations 0 (immediately before expansion and therefore identical to the $N_e = 100$ data), 50, 100, 200, and 400. Results from generation 100,000 come from the final generation of the constant $N_e = 10,000$ population.

A proxy for the homozygous mutation load

Given the evidence that inbreeding depression is caused by homozygosity at numerous partially recessive, deleterious mutations (Charlesworth and Willis 2009), and because deleterious mutations rarely reach frequencies >0.05 in the population (Pritchard 2001), we approximated an individual's overall load of homozygous recessive/partially recessive deleterious mutations by defining mutations as the set of all variants not included in the SNP data (*i.e.*, all variants with MAF < 0.05 in the original sequence data) and summed all such mutations that were homozygous for each individual. We call this measure the "homozygous mutation load". We computed the homozygous mutational load using rare neutral alleles for reasons of computational efficiency, but rare neutral alleles approximated well the behavior of mildly deleterious, partially recessive alleles. For all population sizes, the observed distribution of allele frequencies of simulated neutral alleles with MAF < 0.05 was very similar to the theoretically expected frequency distributions of either partially recessive ($h = 0.25$), very mildly deleterious mutations ($s \sim 1/N_e$) or fully recessive mutations ($h = 0$) of much larger effect ($s \sim 50/N_e$), where h is the dominance coefficient and s the selective coefficient (see supporting information, Figure S1) (Crow and Kimura 1970). Thus, the homozygous mutation load investigated here simulates what would be observed if inbreeding depression is due to the aggregated effect of homozygosity at a large number of recessive to partially recessive, mildly deleterious alleles.

Real SNP data

To compare results from simulated SNP data with results from real human SNP data, we used 1000 randomly selected control individuals of Caucasian descent from the publicly available Molecular Genetics of Schizophrenia–GAIN (Genetic Association Information Network) sample (O'Donovan *et al.* 2008) genotyped on the Affy 6.0 platform. All SNPs passed rigorous quality control metrics (MAF > 0.05, missingness <0.02, Hardy–Weinberg equilibrium *P*-values <0.0001) and individuals who had a missingness rate >0.02 or who were outliers on the first two principal component dimensions derived from an identical-by-state matrix were dropped. Genome-wide, 546,882 SNPs (~1 SNP per 5.1 kb) of 906,600 passed these quality control thresholds. We then selected the first 100 Mb of SNP data from chromosomes 1–10 (after removing the centromere and surrounding heterochromatin if applicable) to create 10 different SNP samples. We derived 10 replicates of the three genomic estimates of F (defined below) from these samples.

Estimates of F

In some contexts, F can be conceptualized as a parameter rather than as an estimate. For example, F_{ped} is a known property given a pedigree and a base population and can be called a parameter of an individual in this context. In the present context, however, F_{ped} and all other F statistics

investigated here are conceptualized as imperfect estimates of the actual level of autozygosity in an individual's genome. The average F estimate in a population depends on the base population (defined as the ancestral population when $F = 0$) and increases as one considers older base populations. However, the average F is not very important in the context considered here. Variation in F among the individuals in the population is required to detect an inbreeding effect, so it is variation in F that we focus on. We use a recent base population—five generations ago—for pedigree inbreeding and the current population for genomic estimates of F as presented in Powell *et al.* (2010).

In each simulated sample, we calculated four alternative estimates of F, as well as two additional estimates of F_{roh} based on shorter and longer megabase thresholds. In the real SNP data sample, we calculated the three genomic estimates of F. To compare the spread of the estimates of F, the variances of F were derived across the 1000 individuals in a sample, and the log (base 10) of these variances was taken to aid interpretability. As there were 10 replicate samples derived from each population size (see above), we then took the mean of the 10 log (base 10) variances and found the standard error around each of these means. Each estimate of F is described below.

1. F_{ped}: F from pedigree inbreeding going back fvie generations as figured from Wright's path formula,

$$F_{\text{ped}_j} = \sum_{i=1:c} \left(\frac{1}{2}\right)^{(n_{ij}+m_{ij}+1)},$$

where m_{ii} and n_{ii} refer to the number of paternal and maternal paths from the ith common ancestor and c refers to the number of common ancestors of individual j. For example, $n_{ij} = m_{ij} = 2$ for cousin–cousin inbreeding, and $c = 2$ when the cousins share two grandparents ($F_{\text{ped}} = 0.0625$) and $c = 1$ when the cousins share one grandparent ($F_{\text{ped}} = 0.03125$). Virtually all variation in F_{ped} is captured within the most recent five generations (see *Results*) and so paths m_{ij} and n_{ij} were not extended beyond this.

2. F_{h} is the canonical estimate of genomic F based on excess SNP homozygosity,

$$F_{\text{h}_j} = \frac{O(H_j) - E(H)}{m - E(H)},$$

where $O(H_j)$ is the observed homozygosity across all SNPs for person j, $E(H) = \sum_i 1 - 2p_i(1-p_i)$ is the expected homozygosity for all people in the sample, and p_i is the MAF for SNPs $i = 1, \ldots, m$. This estimate can be obtained from PLINK (Purcell *et al.* 2007), using the –het command.

3. F_{alt} is an alternative estimate of F predicted to have lower error (Yang *et al.* 2010a):

$$F_{\text{alt}_j} = 1 - \frac{\sum_i \delta_i}{m}.$$

$\delta_i = 1/p_i$ and $1/q_i$ for a homozygote for the minor and the major allele, respectively, and 0 if heterozygote at SNP i, and where q_i is the frequency of the major allele at the ith SNP and $p_i = 1-q_i$.

4. F_{roh} refers to the proportion of the genome (0–1) that is in runs of homozygosity (ROHs):

$$F_{\text{roh}_j} = \frac{\sum_k \text{length}(\text{ROH}_k)}{L}.$$

ROH_k is the kth ROH in individual j's genome, and $L = 10^6$ bp, the length of the genome in both the simulated and the real SNP data sets. ROHs were found using PLINK and defined as stretches of continuously homozygous SNPs spanning at least 1.5 Mb ($\sim$1.65 cM). The lengths of ROH segments generated by a single path should follow an exponential distribution with mean $1/2g$ M, where g is the number of generations since the last common ancestor (Fisher 1954). Thus, the expected length of an autozygous ROH segment caused by a common ancestor $g = 30$ generations in the past is 1.65 cM ($\sim$1.5 Mb).

To compare alternative F_{roh} thresholds, we defined $F_{\text{roh,short}}$ as a run of homozygous SNPs >0.5 Mb and $F_{\text{roh,long}}$ as a run of homozygous SNPs >5 Mb. These three threshold values (0.5 Mb, 1.5 Mb, and 5 Mb) are the same as those used by McQuillan *et al.* (2008) in their investigations into F_{roh} in European populations. $F_{\text{roh,long}}$ is likely to lack many autozygous stretches of ancient origin (*e.g.*, from >10 generations back, which have expected sizes of <5 Mb), but is also likely to be composed of very few ROHs that are not truly autozygous. On the other hand, $F_{\text{roh,short}}$ can detect more ancient autozygous segments (*e.g.*, from $\sim$100 generations back) at the expense of a higher false detection rate—*i.e.*, detecting a higher proportion of ROHs that are allozygous rather than autozygous. Such allozygous haplotypes may be similar superficially, creating runs of homozygosity at measured SNPs, but heterozygous at unmeasured, rare mutations. Using the expected exponential distribution of ROH segments, our thresholds of 0.5 Mb, 1.5 Mb, and 5 Mb should capture 58%, 20%, and 0.5% (respectively) of all ROHs created from common ancestors 50 generations in the past.

Results

Time to most recent common ancestor between spouses

Given that ancient inbreeding is central to the present investigation and that the relationship between mates determines the inbreeding coefficient of their progeny, it is useful to have an intuition of how long ago the most recent common ancestor existed between two mates in a randomly
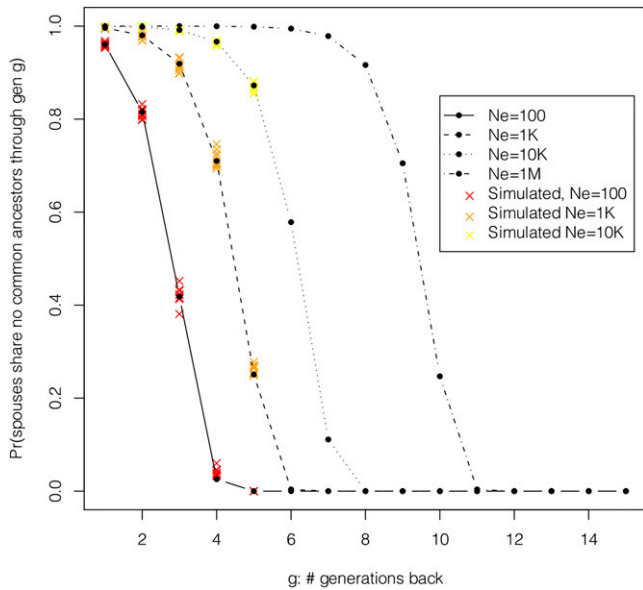
Figure 2 Shown is the probability that mates share no common ancestors in the most recent g generations as a function of population size (see text). The x's are the same values, derived empirically from simulations, for up to 5 generations in the past, and show good agreement with the expected probabilities. Even in large, randomly breeding populations (*e.g.*, 1 million), it is almost certain that at least one ancestor exists in common between two pedigrees within 11 generations.

mating population. A similar issue has been investigated with respect to the amount of overlap in pedigrees between two individuals (Ohno 1996; Derrida *et al.* 2000), as we discuss below. Similarly, Chang (1999) assessed the number of generations one must travel back before reaching the most recent common ancestor of *every* human alive today, and it is surprisingly recent: about 110 generations (∼3100 years) ago (see also Rohde *et al.* 2004; Lachance 2009). Clearly, the most recent common ancestor between just two randomly chosen individuals (*e.g.*, spouses) must be much more recent still. This can be quantified simply in a Fisher–Wright model.

For a population of constant size, consider two randomly chosen individuals (spouses), X and Y. Take a single ancestor of spouse X who lived t generations ago. Given nonoverlapping generations, the probability that this ancestor is the same person as any of spouse Y's ancestors who lived t generations ago is A_t/N_t, where A_t is the total number of ancestors of Y at generation t, and N_t is the total number of all ancestors in the population at generation t. This follows from the fact that the probabilities that the ancestor of X is related to each of Y's ancestors are independent of each other (*i.e.*, multiple ancestors of Y can be the same person). The probability that the ancestor of X is unrelated to any ancestor of Y is therefore $1-A_t/N_t$. Assuming that X also has A_t distinct ancestors at time t, the probability that no ancestor of X is related to any ancestor of Y at time t is $(1-A_t/N_t)^{A_t}$. Finally, the probability, ψ_g, that a randomly chosen spousal pair share no common ancestors (no pedi-

gree inbreeding) up to and including generation g is $\psi_g = \prod_{t=1,\ldots,g}(1-A_t/N_t)^{A_t}$.

The above calculation for ψ_g relies on estimations of A_t and N_t that are not straightforward to derive. As an approximation, $A_t \cong 2^t$ and $N_t \cong N_e$. While the latter assumption is probably close to the truth on average, the former ($A_t \cong 2^t$) must overestimate the number of ancestors at time t because $A_t \leq N_t$ even though A_t grows exponentially while N_t remains constant. Put another way, as one travels back in a pedigree, duplicate ancestors grow increasingly common due to ancient inbreeding. This biases ψ_g to be lower than it should be, but further investigation (not shown) indicates that this bias is minimal, not changing how far one must go back to find a common ancestor by more than a single generation.

Figure 2 shows expected values for ψ_g for four different population sizes: the three investigated in the present simulation as well as for $N_e = 1$ M, which is probably a larger mating pool than most modern humans belong to. We checked our theoretical predictions against simulation results for the three population sizes ($N_e = 100$, 1000, and 10,000) and the five generations of ancestry we recorded. The expected probabilities (solid dots) agree nicely with the observed proportions from the simulated data (colored x's).

Figure 2 shows that mates are likely to share a common ancestor in their recent ancestral past. For populations of size $N_e = 1000$, almost all spousal pairs have at least one common ancestor in common from ≤6 generations in the past. For $N_e = 10,000$, spousal pairs share a common ancestor within 8 generations. Indeed, for every 10-fold increase in population size, the number of generations one must go back before a common ancestor becomes certain increases by <2 generations. Thus, at $N_e = 1,000,000$, spouses are almost certain to share a common ancestor within the last 11 generations. Increasing this population size to 100 million pushes this estimate back only to ∼15 generations. ROHs caused by inbreeding from common ancestors 15 generations in the past are ∼3 Mb in length on average, although, due to the finite number of recombination events per generation, inbreeding events from 15 generations in the past often cause no ROHs. These results are consonant with the results of Derrida *et al.* (2000), who found that there is substantial overlap between pedigree trees of two randomly chosen individuals in a population of size ≤10,000 within 10 generations and that this overlap is complete (both individuals share the exact same ancestors) within ∼20 generations.

Variance of F_{ped}

Given that the contribution to F_{ped} at generation t is independent of the contribution to F_{ped} from previous generations, the variation of F_{ped} for a given population size can be apportioned into the amount contributed by each ancestral generation. Figure 3 shows that the contribution to the variation of F_{ped} decreases log-linearly as a function of the contributing ancestral generation. The bend (nonlinearity) at generation 1 for the samples drawn from the $N_e = 10,000$
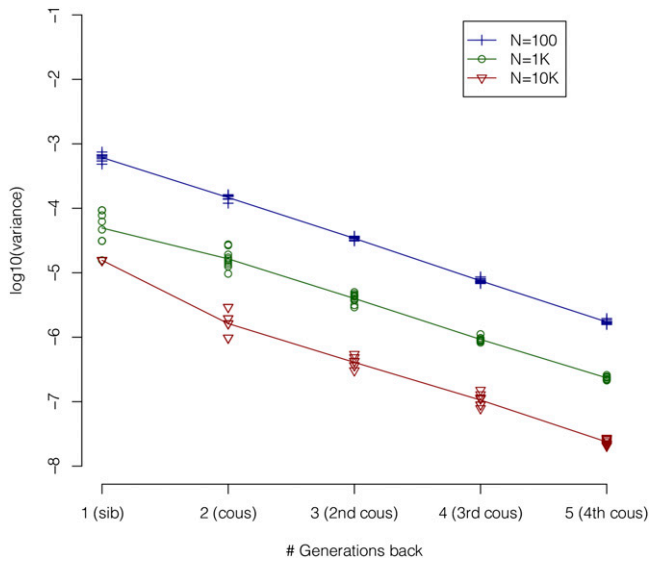
Figure 3 Contribution to var(F_{ped}) from previous generations. Most of the variance in F_{ped} at all population sizes is attributable to recent inbreeding. For example, the variance in F_{ped} due to spouses sharing common ancestors five generations in the past makes up only ~0.2% of the total variance in F_{ped} at all population sizes.

population occurs because several samples had no sib–sib inbreeding, making the $\log_{10}(\text{var}(F_{ped}$ due to sib–sib inbreeding)) $= -\infty$. Such values were set to "missing", biasing the remaining nonmissing values upward.

Our results show that very little variation in F_{ped} was missed by ignoring inbreeding resulting from common ancestors more than five generations back (see *Methods*). This is despite the fact that quite long (*e.g.*, ~10 Mb) ROHs can be created from such inbreeding. Although we had no way of measuring the proportion of variation in genomic estimates of F that was due to each ancestral generation, our results below demonstrate that genomic estimates of F retain considerable variation in large populations after variation in F_{ped} has diminished to ~0.

Variation and error in genomic and pedigree estimates of F

One important aim of the current project was to investigate the effects of population size on the variance of the four estimates of F. Variances were taken across estimates of F for each individual in samples of size 1000 drawn from the three population sizes ($N_e = 100$, 1000, and 10,000). Figure 4 shows the mean (± 1 SE) of the $\log_{10}$ variances of F as a function of N_e. The variance decreases for all estimates of F as N_e increases, caused by the lower levels of ancient and recent inbreeding in larger populations. As expected (Figure 3), F_{ped} shows the least variation at every population size because it does not capture the variation in F caused by the random nature of recombination and segregation. The two estimates of F based on SNP-by-SNP homozygosity (F_h and F_{alt}) show the highest variation, while F_{roh} is intermediate. Figure 4 also shows the variance values for the three geno-

mic estimates of F in 10 replicates of 100 Mb of real SNP data from an unselected (outbred) Caucasian control sample. The variances of F_h and F_{alt} are slightly higher and variance of F_{roh} is slightly lower in the real data than in the $N_e = 10,000$ simulation.

The prediction error variance (PEV) of the three genomic estimates of F is shown in Figure 5. PEV $= \text{var}(F_1 - F_2)/2$, where F_1 is the estimate of F from a random half of SNPs and F_2 is the estimate of F from the other half of SNPs (Powell *et al.* 2010; Yang *et al.* 2010a). PEV provides insight into the amount of error in the genomic estimates of F that exists due to the finite sampling of SNPs; as SNPs become more dense and linkage disequilibrium increases between markers (as occurs with larger N_e), the PEV should decrease. The PEV for all estimates of F is small (<5%) compared to the variances of F shown in Figure 4. As expected, the PEV of every estimate of F also decreases with population size, but the PEV for F_{roh} decreases the most rapidly. Error in F_{roh} decreases more rapidly as a function of N_e because, with dense SNP data (such as exist when $N_e = 10,000$), ROHs are made up of a large number of homozygous SNPs in a row (*e.g.*, a ROH of 1.5 Mb contains ~2000 SNPs in the $N_e = 10,000$ data), and such long sets of markers pick up the same sets of ROHs with very high precision.

Relationships between genomic estimates of F and F_{ped}

Figure 6 shows the correlation coefficients between the three genomic estimates of F and F_{ped}. In all cases, F_{roh} correlates most highly with F_{ped}, reflecting the fact that F_{roh} tends to capture more recent inbreeding than the other two genomic estimates of F. The increasingly large standard errors of the correlations are due to the higher standard errors and lower variances of F_{ped} at larger population sizes.

Finding the optimal measure for detecting inbreeding depression: Relationships between estimates of F and the homozygous mutation load

To understand which estimate of F is likely to be optimal for detecting inbreeding depression effects, we correlated each of the four estimates of F with the homozygous mutation load, which, as explained above, is the leading contender for why inbreeding depression occurs. Figure 7 shows these results. F_{ped} is the worst predictor of the homozygous mutation load, and its disadvantage grows as the population size increases. This result is due to the low variance in F_{ped} in randomly breeding populations and suggests that F_{ped} is likely to be a useful estimate only in samples selected on the basis of recent consanguinity. F_{alt} outperformed F_h because F_{alt} gives more weight to rare variants and because F_{alt} has lower error variance (Figure 5). As predicted, F_{roh} was the most associated estimate with the recessive mutation load at every population size, and this advantage grows as the population size increases (Figure 7).

The superiority of F_{roh} in detecting the homozygous mutation load is even more stark if one considers the unique effects of each estimate of F after controlling for their
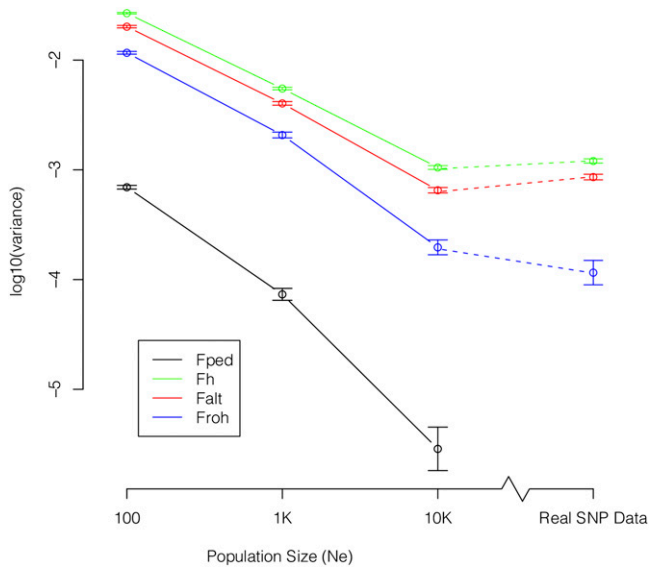
Figure 4 Variance of F (± 1 SE) as a function of N_e. The variance of F_{ped} is the lowest at all population sizes and the variance of F_{roh} is intermediate. Comparing the variance of genomic measures of F in simulated data to the equivalent variances in real SNP data (from an outbred Caucasian sample) suggests that the effective population size of Caucasians is $\sim$10,000 with respect to these statistics.

Figure 5 Prediction error variance (PEV) of genomic estimates of F as a function of N_e. PEV decreases as a function of population size for all genomic estimates of F, but does so most rapidly for F_{roh}.

correlations with F_{roh}. After controlling for F_{roh}, the mean partial correlations (across N_e) between the homozygous mutation load and F_h ($\bar{r}_p = 0.07$), F_{alt} ($\bar{r}_p = 0.15$), and F_{ped} ($\bar{r}_p = 0.09$) are all low, but the reverse is not true: Partial correlations between the homozygous mutation load and F_{roh} remain high after controlling for F_h ($\bar{r}_p = 0.57$), F_{alt} ($\bar{r}_p = 0.50$), and F_{ped} ($\bar{r}_p = 0.68$). This shows that F_{roh} contains much additional information on the homozygous mutation load that is not captured by the other three estimates of F, whereas the other three estimates of F contain little additional information on the homozygous mutation load over what already exists in F_{roh}. The superiority of F_{roh} occurs because a given rare variant is typically homozygous only when the haplotypic segment on which it exists meets another IBD segment within an individual—creating a run of homozygosity. If inbreeding depression is caused by homozygosity at rare mutations, as recent evidence indicates, these results suggest that F_{roh} is likely to be the optimal estimate for detecting it, regardless of the level of inbreeding in the population.

Figure 8 shows a comparison of the correlation coefficients between the homozygous mutation load and F_{roh} calculated from short (>0.5 Mb), moderate (>1.5 Mb), and long (>5 Mb) ROH thresholds, as a function of N_e. $F_{roh,short}$ correlates poorly with the homozygous mutation load when N_e is small. In large populations where inbreeding is more ancient, the situation reverses: F_{roh} defined by short ROHs outperforms the other two F_{roh} estimates. The improvement of $F_{roh,short}$ as a function of population size is due to two factors: (a) denser SNP data in the larger populations allows even short (0.5 Mb) ROHs to discriminate autozygosity ac-
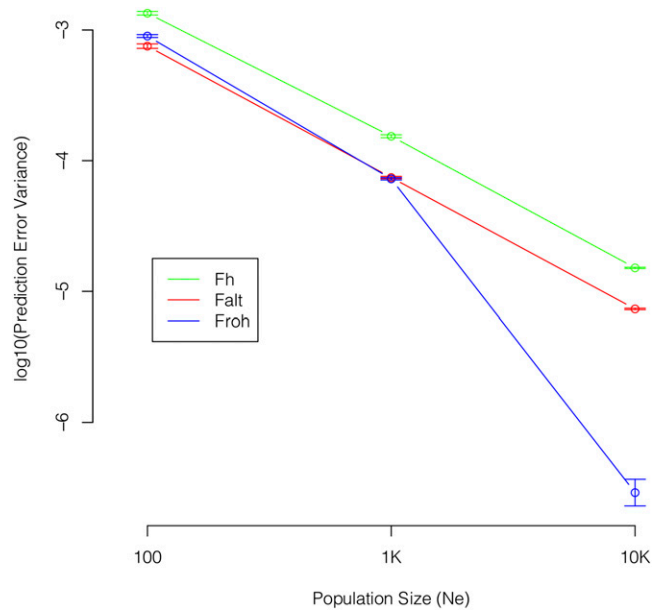
curately and (b) a higher ratio of short to long ROHs in large populations. $F_{roh,long}$ performs poorly when $N_e = 10,000$ due to low variance. F_{roh} defined by the moderate (1.5 Mb) threshold performed intermediately at each population size, providing justification for using this threshold in the current study as it allowed more meaningful comparisons across the three population sizes.

Changes in F following population expansion

We wanted to understand how recent population expansion, such as what occurred in human lineages over the last $\sim$10,000 years ($\sim$350 generations), might alter our conclusions. Figure 9 shows the change in the mean log10(variance) of the four estimates of F as a function of generations since a 100-fold expansion in population size, from $N_e = 100$ to $N = 10,000$. While the variance in F_{ped} decreases immediately (a 97% drop) and remains unchanged thereafter, the three genomic estimates of F change much more gradually. The initial declines in the variances of the three genomic estimates of F are caused by the drop in recent inbreeding in larger populations. However, the declines in variances become much slower thereafter, reflecting the slow decay of linkage disequilibrium and the gradual increase in the number of new segregating variants following a population expansion. For example, after expanding from $N_e = 100$ to $N_e = 10,000$, the standard deviation of F_{roh} decreases by $\sim$70%, from $\sim$0.115 to $\sim$0.035, in the first 100 generations, but decreases only an additional 10% (to $\sim$0.021) over the next 300 generations and requires several thousands of generations to reach its equilibrium value of 0.015. Similarly, starting at $\sim$0.142, the standard deviation of F_{alt} decreases to $\sim$0.069 and $\sim$0.055 in generations 100 and 400, respectively. Our results should extrapolate to larger starting and
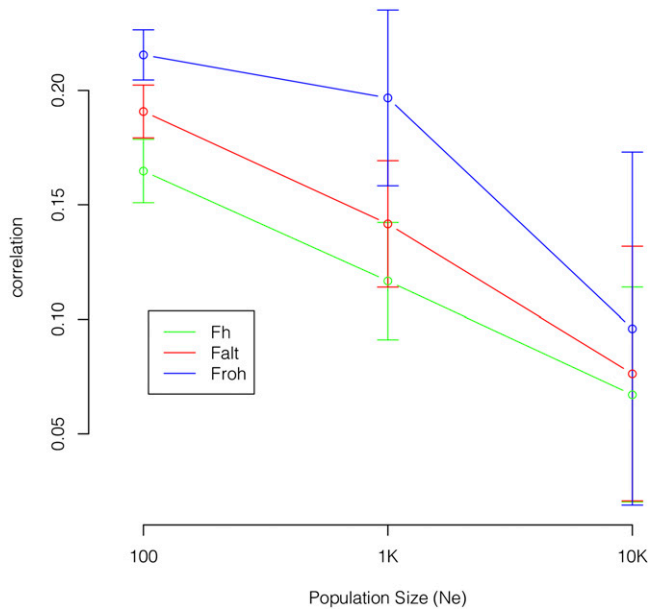
Figure 6 Correlations between F_{ped} and genomic estimates of F as a function of N_e. All correlations between genomic estimates of F and F_{ped} decrease as a function of N_e, but F_{roh} is consistently correlated most highly with F_{ped}.

Figure 7 Correlations between F and homozygous mutation load as a function of N_e. F_{roh} correlates most highly with the homozygous mutation load at all population sizes, and this advantage increases at larger population sizes (where inbreeding becomes more ancient).

ending population sizes (*e.g.*, $N_e = 10,000$ to $N_e = 1,000,000$) and indicate that genomic estimates of F require a considerable time following expansion to reach the low variance values that would be predicted from the larger population sizes.

Figure 10 shows the correlations between estimates of F and the homozygous mutation load following population expansion. This correlation drops considerably for all estimates of F following population expansion, although it remains low for F_{ped}, F_h, and F_{alt} in the 400 generations following the expansion, whereas the correlation with F_{roh} begins to increase after 50 generations. When N_e is small (*i.e.*, 100), rare mutant alleles are tagged by long haplotypes as a result of relatively recent inbreeding. When the population size is increased, these long haplotypes are destroyed by recombination and new ones are not created quickly because new inbreeding occurs slowly. Consequently, the correlation between F and homozygous mutation load declines as did the variance of F (Figure 9). However, as the population evolves at size $N_e = 10,000$ for an increasing number of generations, new mutations arise that exist on only a single haplotype, and the pairings of these, creating homozygosity at rare mutations, are increasingly detected by F_{roh} and eventually by F_h and F_{alt} as a new equilibrium is reached. These results provide further evidence that F_{roh} is likely to be the optimal estimate for detecting inbreeding depression in populations, such as humans, that have expanded rapidly in population size.

Power to detect inbreeding effects

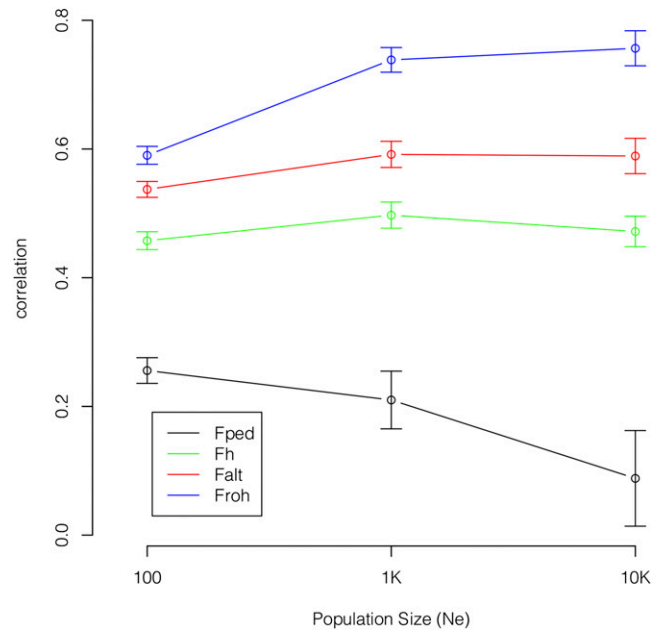The power of detecting inbreeding depression from marker data depends in part on the accuracy of predicting auto-

zygosity with markers, and our results suggest that F_{roh} is optimal for detecting autozygosity arising from both ancient and recent inbreeding. However, power also depends on the variation in F, as well as on the effect size and the sample size. When the level of inbreeding is low, such as in large, randomly mating populations, the variance of F_{roh} is commensurately low (*e.g.*, the standard deviation of $F_{roh} = 0.015$ when $N_e = 10,000$). The standard deviation of F_{roh} in observed human SNP data is smaller yet, ~ 0.012. Given such low variance in the predictor, we were interested in whether inbreeding depression is detectable using F_{roh} in human samples and, if so, what sample sizes would be required as a function of different levels of inbreeding (assayed using different effective population sizes) to achieve adequate power. To do this, we needed an estimate of the likely effect size of inbreeding depression on a complex trait in humans and an estimate of the variance of F_{roh} scaled up to genome-wide levels.

Because the effects of inbreeding on IQ have been investigated more than on any other complex human trait to our knowledge, we used a best estimate of the inbreeding effect on IQ as a guide to the likely inbreeding effect sizes among other human complex traits. Morton (1979) reviewed four large studies of cousin–cousin inbreeding and found that IQ decreased by 0.73 (Schull and Neel 1965) to 0.39 (Kudo *et al.* 1972) points for every 0.01 increase in F_{ped}. Given the standard deviation of IQ (~ 15 points), this translates to a decrease of 0.025–0.05 standard deviations per percentage of inbreeding.

To predict the statistical power of using F_{roh} to detect inbreeding depression in human studies, we also needed
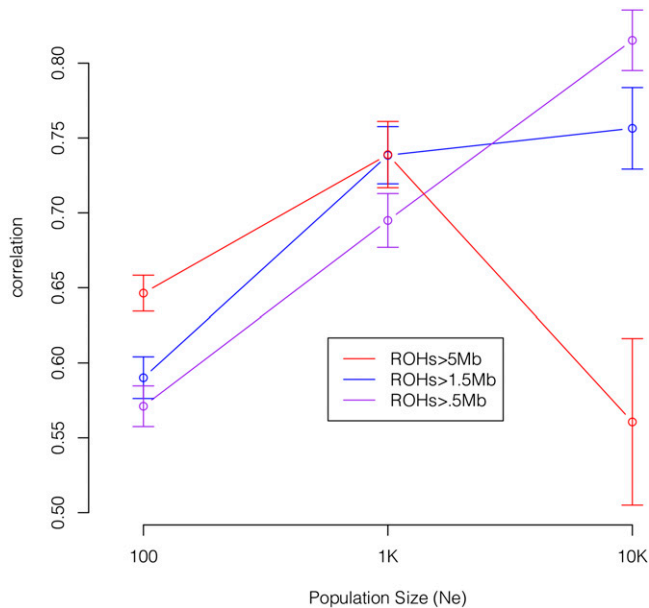
Figure 8 Correlations between alternative thresholds of F_{roh} and the homozygous mutation load as a function of N_e. Runs of homozygosity (ROHs) are defined as stretches of 0.5-Mb, 1.5-Mb, or 5-Mb homozygous SNPs. Long thresholds are optimal for detecting autozygosity in highly inbred populations, whereas shorter thresholds are optimal for detecting autozygosity in outbred populations.

Figure 9 Variance of F (± 1 SE) as a function of generations since population expansion. Whereas F_{ped} drops immediately following a population expansion, the variance in genomic measures of F requires hundreds of generations to reach equilibrium levels.

to estimate what the variance of F_{roh} would be if derived genome-wide rather than across 100 Mb, as simulated in the current study. To do this, we compared the average variance of F_{roh} across 10 replicates of 100 Mb of real SNP data (1.5×10^{-4}) to the genome-wide variance of F_{roh} in the real SNP data (1.79×10^{-5}), an 8.45-fold drop in variance. We thus scaled the variances of F_{roh} down 8.45-fold to account for the decreased variance that would be observed in F_{roh} genome-wide. This led to predicted genome-wide variances (standard deviations) of F_{roh} of 1.4×10^{-3} (0.037) for $N_e =$ 100, 2.5×10^{-4} (0.016) for $N_e = 1000$, 2.63×10^{-5} (0.005) for $N_e = 10{,}000$, and 1.79×10^{-5} (0.004) for real SNP data.

Figure 11 shows the statistical power as a function of sample size of a regression of F_{roh} on an outcome variable assuming a slope of -0.73 (solid lines) and -0.39 (dashed lines), a standard deviation of the outcome variable of 15, and standard deviations of F_{roh} depending on the population size, as detailed in the preceding paragraph. We assume that $\hat{\beta} \sim t(n-2)$ with mean -0.73 or -0.39 and $\text{var}(\hat{\beta}) = (\text{var}(IQ) - \hat{\beta}^2(\text{var}(F_{roh})))/\text{var}(F_{roh})(n-2)$. Our results suggest that sample sizes of <700 are sufficient for achieving adequate (80%) statistical power when the rate of inbreeding is high (similar to that found in a randomly breeding population of size $N_e = 100$), such as could occur in samples selected on the basis of likely recent inbreeding or in small, isolated populations. Our results also suggest that inbreeding depression effects can be detected using F_{roh} in large, ostensibly outbred populations, such as those found in modern industrialized societies, but that large samples (on the order of 12,000–65,000, depending on effect size)
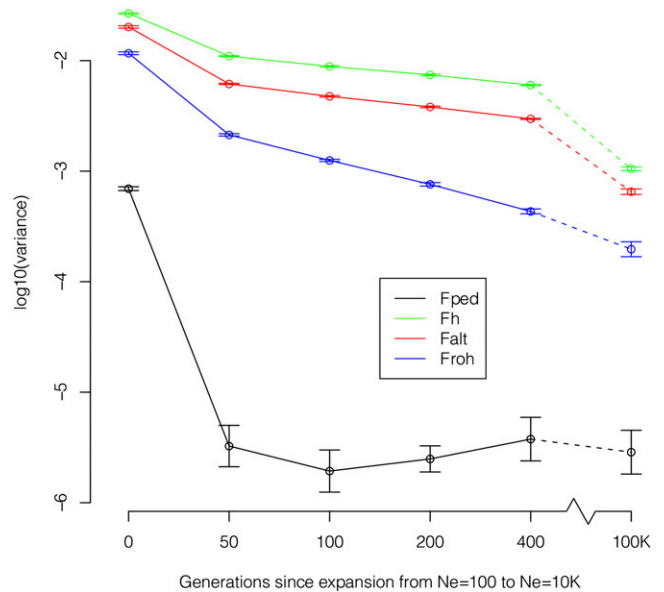
are probably necessary to detect them reliably. It is possible that phenotypes more related to fitness than is IQ would show a larger inbreeding depression effect and thus would require smaller sample sizes than those predicted here to achieve adequate power.

Discussion

Hundreds of scientific investigations have been conducted on the effects of inbreeding since Darwin (1868, 1876) first studied the topic. The inbreeding coefficient, F, has traditionally been defined in terms of pedigree inbreeding (F_{ped}). However, low variation in F_{ped} seriously hampers the ability to detect inbreeding effects in samples that are not selected on the basis of known recent inbreeding. In essence, F_{ped} is a genome-wide expectation for the percentage of the genome that is autozygous, but there is a great deal of variation in autozygosity around this expected percentage caused by the finite number of recombination events per generation. The resulting identical haplotypes can persist in the population for many generations, coming together in offspring of distantly related individuals to create increased levels of homozygosity. The availability of genomic marker panels has made possible new, genomic alternatives to F_{ped} that can better detect such identical haplotypes and therefore even quite ancient inbreeding.

In the last 10 years, a large number of human (reviewed in Ku *et al.* 2010) and animal (reviewed in Chapman *et al.* 2009) studies have investigated the relationship between genomic estimates of F and disease- or fitness-related traits. In general, animal studies have focused on excess homozygosity on a marker-by-marker basis (F_h) whereas human
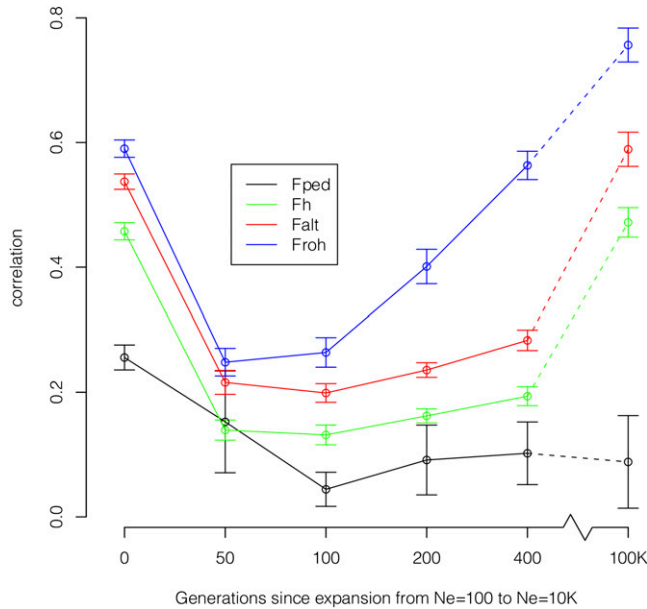
Figure 10 Correlations between F and homozygous mutation load as a function of generations since population expansion. The relationship between the homozygous mutation load and F_{roh} increases the most quickly following a population expansion.

Figure 11 Estimated power to detect inbreeding effects on a human complex trait using F_{roh}. Higher levels of real inbreeding (smaller N_e) lead to higher variance in F_{roh} and thus greater statistical power to detect an inbreeding effect. Large (solid lines) and small (dashed lines) inbreeding effect sizes were derived from a review on the effects of consanguinity on IQ (see text). Arrows show predicted sample sizes required to achieve ~80% power. When inbreeding is high ($N_e = 100$), sample sizes of ~400 are adequate, but in outbred populations ($N_e = 10,000$ or real SNP data), samples sizes >20,000 may be required.

studies, which have the opportunity to use denser SNP platforms, have focused increasingly but not exclusively on runs of homozygosity (F_{roh}).

In this article, we show that F_{roh} is preferable to F_{ped} and to marker-by-marker estimates of F (F_{alt} and F_h) for detecting both recent and ancient inbreeding, even in cases where the level of inbreeding is likely to be high. F_{roh} correlates most highly with the homozygous mutation load, the putative causal mechanism underlying inbreeding depression, and this advantage is especially pronounced following a population expansion, such as has occurred in recent human history. Moreover, F_{roh} has low prediction error variance, especially when SNP density is high. Our findings provide empirical justification to the growing literature using F_{roh} to study complex traits in humans (Lencz *et al.* 2007; Nalls *et al.* 2009; Spain *et al.* 2009; Vine *et al.* 2009; Enciso-Mora *et al.* 2010; Hosking *et al.* 2010; Yang *et al.* 2010b).

Nevertheless, the variance in F_{roh} in large simulated ($N_e = 10,000$) and observed Caucasian SNP data sets is low, and because of this, there is likely to be little power to detect inbreeding effects in unselected (with respect to recent inbreeding) samples unless samples sizes are large. We estimate that sample sizes between 12,000 and 65,000 would be required to regularly detect previously reported IQ-inbreeding effects using F_{roh} in unselected samples. Thus, current studies investigating the effects of F_{roh} on human complex traits that have samples sizes <3000 and that have failed to find significant inbreeding effects (Nalls *et al.* 2009; Spain *et al.* 2009; Vine *et al.* 2009; Enciso-Mora *et al.* 2010; Hosking *et al.* 2010) are likely to be underpowered. Furthermore, small studies (*e.g.*, $n < 1000$) that do find significant
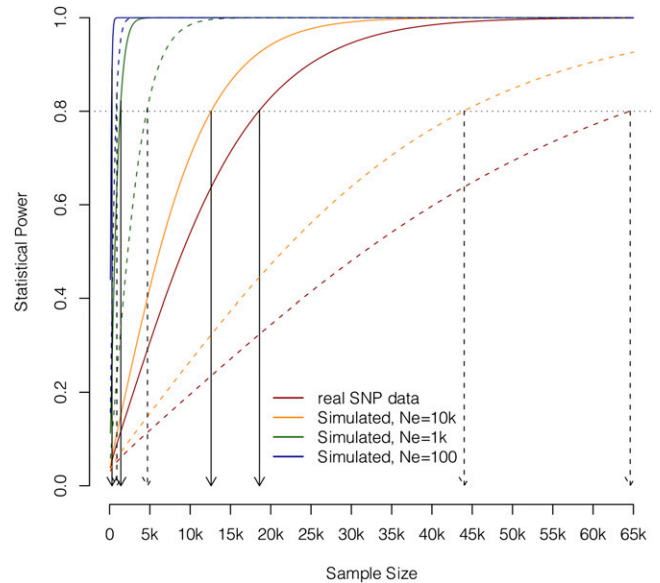
inbreeding depression effects using F_{roh} (*e.g.*, Lencz *et al.* 2007) may greatly overestimate the size of the effects.

Our findings suggest two strategies for achieving sufficient power in studying inbreeding depression of complex traits using a genomic estimate of F. The first is to conduct analyses on very large (*e.g.*, $n > 10,000$) samples, such as those being put together by multisite consortia (International Schizophrenia Consortium 2009). The second is to conduct analyses on smaller samples (perhaps ~1000–3000) from populations where the variation in inbreeding is likely to be high, such as in population isolates (Rudan *et al.* 2009) or in cultures where close inbreeding is common (Bittles and Black 2010a). In either case, F_{roh} is likely to be a more powerful approach for detecting inbreeding depression than any other alternative.

While much of the focus in this article has been on understanding the behavior of estimates of F in human populations, our general results should apply to nonhuman animal populations as well. Animals with larger effective population sizes than humans will manifest lower levels of variation in F than those reported here and will require commensurately larger sample sizes to detect inbreeding effects in nonselected samples. On the other hand, many species have smaller effective population sizes than humans. Moreover, fitness traits in other species may show larger inbreeding depression effects than the human example explored here and may be detectable with smaller sample sizes.

The simulation parameters explored in this article were chosen to mimic the human genome. Recombination rates and recombination hotspot rates can vary at least threefold across mammals (Jensen-Seaman *et al.* 2004; Kauppi *et al.* 2004). Nevertheless, the basic conclusions made here are not overly dependent on the specific parameters chosen and can readily be applied to animals with different recombination parameters. When distance is conceptualized in centimorgans rather than base pair units, different recombination rates lengthen or shorten the genome vis-à-vis our simulation. This would slightly lower or raise (respectively) the variance of F_{roh}, as shown above when variation in F_{roh} was extrapolated to its expected genome-wide value. Similarly, when genomic distance is measured in centimorgans, different recombination hotspot parameters change the variation in SNP density vis-à-vis our simulation. This should have minimal effect on the ROHs detected, and therefore on F_{roh}, because even the shortest ROH encompasses $\sim$175 hotspots on average. Thus, while the specific values in our results depend on the specific simulation parameters used, and are therefore most relevant to human populations, the qualitative conclusions regarding the advantages of F_{roh} over the alternative estimates of F apply across a wide range of parameters and extend to nonhuman animals as well.

Inbreeding has had a central place in the field of population genetics since its inception. Its importance derives not only from its relevance to many theoretical concepts in population genetics. Evolutionists have used inbreeding as a way to gauge which traits are likely to have been under ancestral selection given that traits most affected by it tend to be most related to fitness (Roff 1997). At a more practical level, inbreeding can have important health consequences in human populations (Bittles and Black 2010b; Rudan *et al.* 2003a). Rates of recent inbreeding are not low across the world. For example, progeny from second-cousin or closer marriages ($F_{ped} > 0.016$) are estimated to account for 10.4% of the global human population. In this article, we have shown that rates of inbreeding due to distant common ancestors in large outbred populations are high enough to have detectable effects, and we have demonstrated the optimal way of identifying such ancient inbreeding. Investigations into ancient inbreeding effects should help investigators understand the evolutionary forces acting on the genes underlying trait variation and whether inbreeding represents an important risk factor in disease.

Acknowledgments

Literature Cited

Abaskuliev, A. A., and G. V. Skoblo, 1975 Inbreeding, endogamy and exogamy among relatives of schizophrenia patients. Genetika 11: 145–148.

Afzal, M., 1988 Consequences of consanguinity on cognitive behavior. Behav. Genet. 18: 583–594.

Bittles, A. H., and M. L. Black, 2010a Evolution in health and medicine Sackler colloquium: consanguinity, human evolution, and complex diseases. Proc. Natl. Acad. Sci. USA 107(Suppl. 1): 1779–1786.

Bittles, A. H., and M. L. Black, 2010b The impact of consanguinity on neonatal and infant health. Early Hum. Dev. 86: 1779–1786.

Brown, D. E., 1991 *Human Universals*. McGraw-Hill, New York.

Bulayeva, K. B., S. M. Leal, T. A. Pavlova, R. M. Kurbanov, S. J. Glatt *et al.*, 2005 Mapping genes of complex psychiatric diseases in Daghestan genetic isolates. Am. J. Med. Genet. Part B Neuropsych. Genet. 132: 76–84.

Carothers, A. D., I. Rudan, I. Kolcic, O. Polasek, C. Hayward *et al.*, 2006 Estimating human inbreeding coefficients: comparison of genealogical and marker heterozygosity approaches. Ann. Hum. Genet. 70: 666–676.

Chadeau-Hyam, M., C. J. Hoggart, P. F. O'Reilly, J. C. Whittaker, M. De Iorio *et al.*, 2008 Fregene: simulation of realistic sequence-level data in populations and ascertained samples. BMC Bioinform. 9: 364.

Chang, J. T., 1999 Recent common ancestors of all present-day individuals. Adv. Appl. Probab. 4: 1002–1026.

Chapman, J. R., S. Nakagawa, D. W. Coltman, J. Slate, and B. C. Sheldon, 2009 A quantitative review of heterozygosity-fitness correlations in animal populations. Mol. Ecol. 18: 2746–2765.

Charlesworth, B., and D. Charlesworth, 1999 The genetic basis of inbreeding depression. Genet. Res. 74: 329–340.

Charlesworth, D., and J. H. Willis, 2009 The genetics of inbreeding depression. Nat. Rev. Genet. 10: 783–796.

Crow, J. F., and M. Kimura, 1970 *An Introduction to Population Genetics Theory*. Harper & Row, New York.

Darwin, C., 1868 *The Variation of Animals and Plants Under Domestication*. Appleton, New York.

Darwin, C., 1876 *The Effects of Crossing and Self Fertilization in the Vegetable Kingdom*. John Murray, London.

DeRose, M. A., and D. A. Roff, 1999 A comparison of inbreeding depression in life-history and morphological traits in animals. Evolution 53: 1288–1292.

Derrida, B., S. C. Manrubia, and D. H. Zanette, 2000 On the genealogy of a population of biparental individuals. J. Theor. Biol. 203: 303–315.

Enciso-Mora, V., F. J. Hosking, and R. S. Houlston, 2010 Risk of breast and prostate cancer is not associated with increased homozygosity in outbred populations. Eur. J. Hum. Genet. 18: 909–914.

Fenner, J. N., 2005 Cross-cultural estimation of the human generation interval for use in genetics-based population divergence studies. Am. J. Phys. Anthropol. 128: 415–423.

Fisher, R. A., 1954 A fuller theory of "junctions" in inbreeding. Heredity 8: 187–197.

Franklin, I. R., 1977 The distribution of the proportion of the genome which is homozygous by descent in inbred individuals. Theor. Popul. Biol. 11: 60–80.

Gibson, J., N. E. Morton, and A. Collins, 2006 Extended tracts of homozygosity in outbred human populations. Hum. Mol. Genet. 15: 789–795.

Gindilis, V. M., R. G. Gainullin, and L. M. Shmaonova, 1989 Genetico-demographic patterns of the prevalence of various forms of endogenous psychoses. Genetika 25: 734–743.

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

Hosking, F. J., E. Papaemmanuil, E. Sheridan, S. E. Kinsey, T. Lightfoot et al., 2010 Genome-wide homozygosity signatures and childhood acute lymphoblastic leukemia risk. Blood 115: 4472–4477.

International Schizophrenia Consortium, 2009 A framework for interpreting genome-wide association studies of psychiatric disorders. Mol. Psychiatry 14: 10–17.

Jensen-Seaman, M. I., T. S. Furey, B. A. Payseur, Y. Lu, K. M. Roskin et al., 2004 Comparative recombination rates in the rat, mouse, and human genomes. Genome Res. 14: 528–538.

Kauppi, L., A. J. Jeffreys, and S. Keeney, 2004 Where the crossovers are: recombination distributions in mammals. Nat. Rev. Genet. 5: 413–424.

Ku, C. S., N. Naidoo, S. M. Teo, and Y. Pawitan, 2010 Regions of homozygosity and their impact on complex diseases and traits. Hum. Genet. 129: 1–15.

Kudo, A., K. Ito, and K. Tanaka, 1972 Genetic studies on inbreeding in some Japanese populations. X. The effects of parental consanguinity on psychometric measurements, school performances and school attendance in Shizuoka school-children. Jpn. J. Hum. Genet. 17: 231–248.

Lachance, J., 2009 Inbreeding, pedigree size, and the most recent common ancestor of humanity. J. Theor. Biol. 261: 238–247.

Lebel, R. R., and W. B. Gallagher, 1989 Wisconsin consanguinity studies. II: familial adenocarcinomatosis. Am. J. Med. Genet. 33: 1–6.

Lencz, T., C. Lambert, P. DeRosse, K. E. Burdick, T. V. Morgan et al., 2007 Runs of homozygosity reveal highly penetrant recessive loci in schizophrenia. Proc. Natl. Acad. Sci. USA 104: 19942–19947.

Leutenegger, A. L., B. Prum, E. Genin, C. Verny, A. Lemainque et al., 2003 Estimation of the inbreeding coefficient through use of genomic data. Am. J. Hum. Genet. 73: 516–523.

Li, L. H., S. F. Ho, C. H. Chen, C. Y. Wei, W. C. Wong et al., 2006 Long contiguous stretches of homozygosity in the human genome. Hum. Mutat. 27: 1115–1121.

McQuillan, R., A. L. Leutenegger, R. Abdel-Rahman, C. S. Franklin, M. Pericic et al., 2008 Runs of homozygosity in European populations. Am. J. Hum. Genet. 83: 359–372.

Morton, N. E., 1979 Effect of inbreeding on IQ and mental retardation. Proc. Natl. Acad. Sci. USA 75: 3906–3908.

Nalls, M. A., R. J. Guerreiro, J. Simon-Sanchez, J. T. Bras, B. J. Traynor et al., 2009 Extended tracts of homozygosity identify novel candidate genes associated with late-onset Alzheimer's disease. Neurogenetics 10: 183–190.

O'Donovan, M. C., N. Craddock, N. Norton, H. Williams, T. Peirce et al., 2008 Identification of loci associated with schizophrenia by genome-wide association and follow-up. Nat. Genet. 40: 1053–1055.

Ohno, S., 1996 The Malthusian parameter of ascents: What prevents the exponential increase of one's ancestors? Proc. Natl. Acad. Sci. USA 93: 15276–15278.

Powell, J. E., P. M. Visscher, and M. E. Goddard, 2010 Reconciling the analysis of IBD and IBS in complex trait studies. Nat. Rev. Genet. 11: 800–805.

Pritchard, J. K., 2001 Are rare variants responsible for susceptibility to complex diseases? Am. J. Hum. Genet. 69: 124–137.

Purcell, S., B. Neale, K. Todd-Brown, L. Thomas, M. A. Ferreira et al., 2007 PLINK: a tool set for whole-genome association and population-based linkage analyses. Am. J. Hum. Genet. 81: 559–575.

Pusey, A., and M. Wolf, 1996 Inbreeding avoidance in animals. Trends Ecol. Evol. 11: 201–206.

Roff, D. A., 1997 Evolutionary Quantitative Genetics. Chapman & Hall, New York.

Rohde, D. L., S. Olson, and J. T. Chang, 2004 Modelling the recent common ancestry of all living humans. Nature 431: 562–566.

Rudan, I., D. Rudan, H. Campbell, A. Carothers, A. Wright et al., 2003a Inbreeding and risk of late onset complex disease. J. Med. Genet. 40: 925–932.

Rudan, I., N. Smolej-Narancic, H. Campbell, A. Carothers, A. Wright et al., 2003b Inbreeding and the genetic complexity of human hypertension. Genetics 163: 1011–1021.

Rudan, I., T. Skaric-Juric, N. Smolej-Narancic, B. Janicijevic, D. Rudan et al., 2004 Inbreeding and susceptibility to osteoporosis in Croatian island isolates. Coll. Anthropol. 28: 585–601.

Rudan, I., A. Marusic, S. Jankovic, K. Rotim, M. Boban et al., 2009 "10001 Dalmatians:" Croatia launches its national biobank. Croat. Med. J. 50: 4–6.

Schull, W. J., and J. V. Neel, 1965 The Effects of Inbreeding on Japanese Children. Harper & Row, New York.

Shami, S. A., R. Qaisar, and A. H. Bittles, 1991 Consanguinity and adult morbidity in Pakistan. Lancet 338: 954.

Spain, S. L., J. B. Cazier, R. Houlston, L. Carvajal-Carmona, and I. Tomlinson, 2009 Colorectal cancer risk is not associated with increased levels of homozygosity in a population from the United Kingdom. Cancer Res. 69: 7422–7429.

Stam, P., 1980 The distribution of the fraction of the genome identical by descent in finite random mating populations. Genet. Res. 35: 131–155.

Takahata, N., Y. Satta, and J. Klein, 1995 Divergence time and population size in the lineage leading to modern humans. Theor. Popul. Biol. 48: 198–221.

Tenesa, A., P. Navarro, B. J. Hayes, D. L. Duffy, G. M. Clarke et al., 2007 Recent human effective population size estimated from linkage disequilibrium. Genome Res. 17: 520–526.

Vine, A. E., A. McQuillin, N. J. Bass, A. Pereira, R. Kandaswamy et al., 2009 No evidence for excess runs of homozygosity in bipolar disorder. Psychiatr. Genet. 19: 165–170.

Woods, C. G., J. Cox, K. Springell, D. J. Hampshire, M. D. Mohamed et al., 2006 Quantification of homozygosity in consanguineous individuals with autosomal recessive disease. Am. J. Hum. Genet. 78: 889–896.

Wright, S., 1922 Coefficients of inbreeding and relationships. Am. Nat. 56: 330–339.

Wright, S., 1977 Evolution and the Genetics of Populations, Vol. 3: Experimental Results and Evolutionary Deductions. University of Chicago Press, Chicago.

Yang, J., B. Benyamin, B. P. McEvoy, S. Gordon, A. K. Henders et al., 2010a Common SNPs explain a large proportion of the heritability for human height. Nat. Genet. 42: 565–569.

Yang, T. L., Y. Guo, L. S. Zhang, Q. Tian, H. Yan et al., 2010b Runs of homozygosity identify a recessive locus 12q21.31 for human adult height. J. Clin. Endocrinol. Metab. 95: 3777–3782.

Communicating editor: N. A. Rosenberg

GENETICS

Quantification of Inbreeding Due to Distant Ancestors and Its Detection Using Dense Single Nucleotide Polymorphism Data

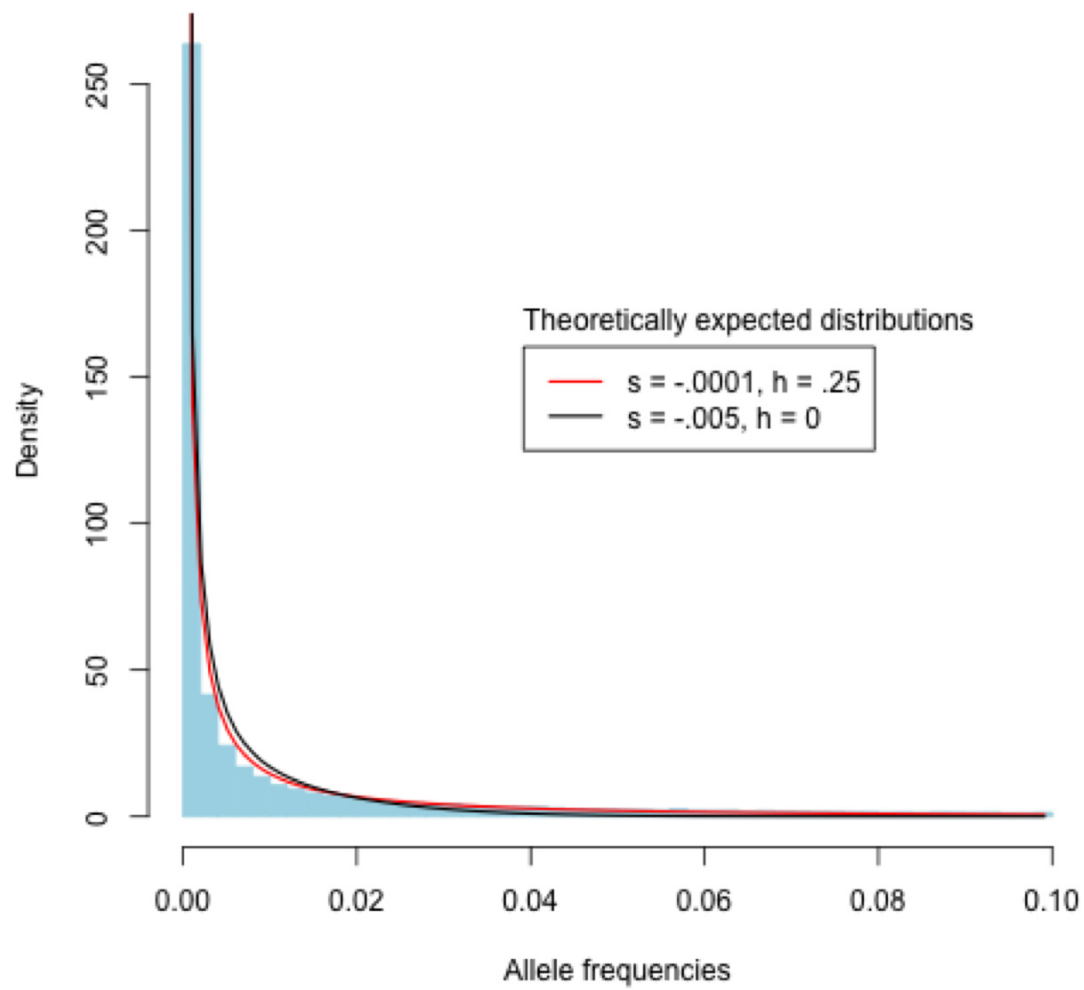Matthew C. Keller, Peter M. Visscher, and Michael E. Goddard

Figure S1 Histogram of observed simulated neutral allele frequencies and theoretically expected distributions (Ne=10K; MAF < .10). The figure shows that our use of neutral alleles is a good approximation for slightly deleterious, partially recessive alleles.