doi:10.1111/j.1558-5646.2010.00966.x

HETEROZYGOSITY-FITNESS CORRELATIONS: A TIME FOR REAPPRAISAL

Marta Szulkin,[1] Nicolas Bierne,[2,3] and Patrice David[4,5]

[1]*Department of Zoology, Edward Grey Institute, University of Oxford, South Parks Road, Oxford OX1 3PS, United Kingdom*

[2]*Université Montpellier 2, Place Eugène Bataillon, F34095 Montpellier, France*

[3]*CNRS, Institut des Sciences de l'Evolution (UMR5554), Station Méditerranéenne de l'Environnement Littoral, 1 Quai de la Daurade, F34200 Sète, France*

[4]*CNRS, Centre d'Ecologie Fonctionnelle et Evolutive (UMR 5175), 1919 Route de Mende, F34293 Montpellier, France*

[5]*E-mail: patrice.david@cefe.cnrs.fr*

Received October 9, 2009
Accepted January 6, 2010

Owing to the remarkable progress of molecular techniques, heterozygosity-fitness correlations (HFCs) have become a popular tool to study the impact of inbreeding in natural populations. However, their underlying mechanisms are often hotly debated. Here we argue that these "debates" rely on verbal arguments with no basis in existing theory and inappropriate statistical testing, and that it is time to reconcile HFC with its historical and theoretical fundaments. We show that available data are quantitatively and qualitatively consistent with inbreeding-based theory. HFC can be used to estimate the impact of inbreeding in populations, although such estimates are bound to be imprecise, especially when inbreeding is weak. Contrary to common belief, linkage disequilibrium is not an alternative to inbreeding, but rather comes with some forms of inbreeding, and is not restricted to closely linked loci. Finally, the contribution of local chromosomal effects to HFC, while predicted by inbreeding theory, is expected to be small, and has rarely if ever proven statistically significant using adequate tests. We provide guidelines to safely interpret and quantify HFCs, and present how HFCs can be used to quantify inbreeding load and unravel the structure of natural populations.

KEY WORDS: **Fitness, inbreeding, population genetics, quantitative genetics.**

Understanding how natural selection acts in contemporary populations is a major goal for evolutionary biologists. Although this task can be undertaken at the phenotypic level using quantitative genetics and selection gradient analysis, the gene level is mainly studied through molecular techniques. However the link between molecular and phenotypic variation is often complex. This link can be dissected based on correlated inheritance patterns of molecular markers and traits in controlled crosses (QTL analysis). Another, less-demanding method is to study the statistical association between molecular genotypes and traits under selection in natural populations. In particular, one kind of such associations, the correlation between individual multilocus heterozygosity (MLH) at allozyme or microsatellite loci and fitness-related traits, also called heterozygosity-fitness correlation (hereafter, HFC), has been studied and discussed for more than three decades. All these studies share a common, fundamental aim, that is, to study natural selection in wild populations, but with different points of view. Some of them ultimately intend to identify selection on one or a few polymorphic genes, whereas others are interested in much more general sources of variation in fitness, such as inbreeding. Therefore this literature is prone to controversy, and it has proven difficult to achieve a unanimous explanation for HFC. Our aim is to contribute to settle these debates and to clarify how HFC can be measured and interpreted, and what it can, and cannot, be used for.

Historically, debates on HFC were concerned with the selective neutrality of allozymes. The focus of the debates has changed with the advent of noncoding DNA markers. The repeated

© 2010 The Author(s). Journal compilation © 2010 The Society for the Study of Evolution.

observation of HFC with such markers (Coltman and Slate 2003; Chapman et al. 2009) calls for a general explanation that would not imply any direct effect of the marker loci on phenotypes. This explanation also has to account for the fact that HFC is not a strong and consistent phenomenon, but rather a weak and unstable signal that shows up from time to time in various organisms in a context-dependent fashion (Britten 1996; Chapman et al. 2009). We argue that such an explanation has existed from the beginning (Ohta 1971, 1973), namely, that heterozygosity at neutral markers is correlated to heterozygosity at both linked and unlinked selected loci through genetic associations. However, the latter only arise in particular contexts, involving a form of inbreeding sensu lato, such as small population size, nonrandom mating, population admixture, or bottlenecks.

Recent HFC literature emphasizes that heterozygosity at a few loci is poorly correlated with inbreeding (i.e., genomic heterozygosity) in large, panmictic populations at equilibrium (Balloux et al. 2004; Pemberton 2004; Slate et al. 2004). Although this assertion is valid, it has been improperly used to dismiss inbreeding as a source of HFC in natural populations. Instead of acknowledging that populations sometimes depart from the preconceived idea of being panmictic, large and at equilibrium, it has been proposed that inbreeding cannot generate HFCs. Instead, HFCs would be caused by nonneutral genes physically linked to molecular markers, generating so-called local effects (e.g., Hansson et al. 2004; Lieutenant-Gosselin and Bernatchez 2006; Charpentier et al. 2008; Malo and Coulson 2009). This idea has often been considered as an alternative explanation to HFCs. We believe that this intense focus on linkage is out of proportion with its true role. We will show that linkage is not an alternative to inbreeding. Indeed, HFCs require correlations in heterozygosity among loci that inevitably imply some level of inbreeding. Although linkage can strengthen these correlations, it does not generate them.

We start by introducing the confusing terminology of HFC, and its historical underpinnings (Section "HFCs: Terminology and Historical Context"). We then outline the theoretical basis of HFC, and the population characteristics that must be met to observe it (Section "The Origin of HFCs: Theory and Practice"). Section "What Meta-Analyses Tell Us about HFC" summarizes evidence for HFCs in a meta-analytical context, followed by a discussion on statistical pitfalls surrounding HFCs and their implications in what is often presented as a local versus general effects debate (Section "Variation in HFC among Loci and the Dichotomy of Local versus General Effects"). Section "HFCs and Alternative Measures of Inbreeding: Future Directions" presents a step-by-step guide on how to derive and interpret robust HFC estimates and, lastly, highlights future directions for the application of HFCs in natural populations.

HFCs: TERMINOLOGY AND HISTORICAL CONTEXT

Inbreeding (terms are defined in Table 1) often reveals spectacular variation in fitness among genotypes that would otherwise remain unexpressed, and has thus fascinated biologists for decades (Darwin 1876; Charlesworth and Charlesworth 1987; Keller and Waller 2002). The terms inbreeding depression and heterosis (see Table 1) were invented early in the twentieth century; these phenomena were first studied in detail in agricultural crops, domestic breeds, and invertebrate model systems (Shull 1908; Gowen 1952). For example, inbred maize lines have a poor agricultural yield compared to their ancestors (inbreeding depression), but the cross-progeny of two different inbred lines often matches or even exceeds the performances of the ancestors (heterosis). The mechanisms behind these two effects were intensely debated (review in Charlesworth and Willis 2009). Because inbreeding and outbreeding modify the heterozygosity at all loci of a genome simultaneously, the mode of action of each individual locus on fitness remained unknown. Both directional dominance and overdominance (Table 1) could explain the observations. Since then, an increasing weight of quantitative genetic evidence points to directional dominance as the main source of inbreeding depression and heterosis (Lynch and Walsh 1998; Charlesworth and Charlesworth 1999; Charlesworth and Willis 2009).

In the decades 1970–2000, the development of molecular tools in population genetics, initiated by Hubby and Lewontin's study of allozymic diversity in *Drosophila* (Hubby and Lewontin 1966; Lewontin and Hubby 1966), made it possible to assess heterozygosity at individual loci, instead of using global measures such as inbreeding coefficients. The number of heterozygous loci in an individual at a given, predefined set of marker loci is called MLH. HFC are statistical associations between MLH and fitness traits. From the 1970s on, HFCs rapidly became of great interest to evolutionary biologists (e.g., Schaal and Levin 1976; Zouros et al. 1980; Mitton and Grant 1984; Britten 1996; David 1998).

Unfortunately, HFC studies use a confusing terminology. During the golden age of allozymes (late 1960s to late 1990s), the debate revolved around whether HFCs reflected direct overdominance at allozyme loci, or some indirect mechanism. Early in the 1970s, many considered HFC, overdominance, and heterosis to be one and the same thing and, consequently, used the terms more or less interchangeably. Later, HFCs were observed again using microsatellites. Most people agree that microsatellites are predominantly noncoding and neutral loci (Jarne and Lagoda 1996). Therefore, their relationship with phenotypes must be indirect. This modifies our perception of historically inherited terminology, which is partly rooted in a neutralist-selectionist debate that does not apply to recent HFC studies. For example, the term associative overdominance (AO, Table 1) remains in use, although it

Table 1. Terminology.

Inbreeding—It occurs when relatives mate, whereby two copies of the same ancestral gene are transmitted to a single zygote. The inbreeding coefficient f refers to the probability that two alleles at a locus in an individual are identical by descent. Because all members of a species are related to some degree, no truly "outbred" individual can exist. Thus, inbreeding, relatedness, and identity by descent are relative rather than absolute measures and have been used in several contexts (Jacquard 1975; Rousset 2002). Fixation indices (F_{is}, F_{ST}, Wright 1922) allow to quantify inbreeding at the population level with respect to nonrandom mating within demes (F_{is}) and population subdivision (F_{ST}). No such measure exists for inbreeding due to small population size or bottlenecks. f can be estimated for individuals whose pedigree is known. The temporal depth of the pedigree sets a limit to such estimates, as the oldest known ancestors are assumed to be unrelated and de facto constitute the reference generation.
Inbreeding depression—A decline in mean fitness (or mean phenotype) in inbred relative to outbred individuals. Both directional dominance and overdominance can produce inbreeding depression.
Inbreeding load—(1) Reduction in population mean fitness as a consequence of inbreeding; or (2) a measure of the slope of the regression of a fitness trait on the inbreeding coefficient f. This slope depends on the abundance and effect of deleterious alleles in genomes—ranging from recessive lethals to very slightly deleterious mutations (Ohta 1992; Kondrashov 1995).
Heterosis (hybrid vigor)—The fitness (or phenotypic) advantage of cross-progeny between two groups (strains, inbred lines or populations), compared to within-group progeny (Shull 1908; Crow 2008). Heterosis can appear in crosses of historically diverged populations or subspecies. Note that depending on the amount of divergence and on the trait studied, heterosis may become negative (a phenomenon referred to as outbreeding depression). Important: Heterosis and overdominance are sometimes used interchangeably—this is a historical legacy from the 60s, when direct overdominance, inferred from allozymic markers, was thought to cause heterosis. Accumulating evidence suggests that directional dominance is in fact largely responsible for both heterosis and inbreeding depression. In the same vein, HFCs have sometimes been called "allozyme-associated heterosis" in the allozyme literature.
Dominance—When genetic effects of two alleles at a locus are not additive, the allele whose homozygous phenotype lies closest to the heterozygote is said to be partially dominant over the other (recessive) allele. Directional dominance occurs when alleles that tend to decrease a phenotype are, on average, mostly recessive. Thus, in heterozygous individuals, deleterious recessive alleles brought in by one parent are concealed by a dominant allele inherited from the other parent. Directional dominance is currently believed to be the main cause for inbreeding depression and heterosis (Lynch and Walsh 1998; Charlesworth and Charlesworth 1999; Crow 2008).
Overdominance—A form of (direct) selection whereby heterozygous individuals exceed the performance of both homozygotes. Dominance and overdominance have different evolutionary consequences: Under directional dominance, inbreeding depression results from a recurrent influx of deleterious mutations, whereas overdominance maintains the polymorphism. Overdominance has now been dismissed as a major cause of inbreeding depression, although it may still play a limited role (Crow 2008).
Associative overdominance (AO, apparent/pseudooverdominance)—Historically introduced by Ohta and Kimura (1969a; 1970), AO is the difference in average fitness between heterozygotes and homozygotes at a neutral locus caused by a statistical association with a (distinct) fitness locus. As such, AO is one of the possible origins of HFC, and indeed the only possible origin in the case of neutral marker loci such as microsatellites. This definition of AO is used throughout our article. Note however that the term AO also appears in relation to the influence of selected loci on the long-term dynamics of allele frequencies at neutral loci (rather than on their correlation with fitness traits). Depending on population and genome characteristics, deleterious alleles may accelerate the fixation of neutral alleles (Charlesworth 1994) or decelerate it (Pamilo and Palsson 1998). The latter case occurs in small populations and genomes with little recombination, and has been called "AO" because it tends to preserve polymorphism longer than expected under neutrality, and therefore mimics an attenuated form of overdominance. This second definition of AO is not used in this article.
Linkage disequilibrium (LD, gametic phase disequilibrium)—Nonrandom association of alleles at two loci in gametes within a population (Lewontin and Kojima 1960; Hedrick 2006a). This can be due to genetic drift, migration, or selection (Hill and Robertson 1968; Sved 1968; Ohta and Kimura 1969b; Weir and Hill 1980). LD is rapidly eroded by recombination. For this reason, it lasts longer when the two loci are tightly physically linked. However, the term is misleading, as physical linkage is not required to generate LD.
Identity disequilibrium (ID)—correlation in heterozygosity and/or homozygosity across loci (Weir and Cockerham 1973). Unlike LD, ID can arise between any two loci in an infinite population without migration or selection, provided some consanguineous matings occur (Bennett and Binet 1956). Under random mating, drift, bottlenecks, and admixture can create ID: in these cases LD co-occurs with ID because when alleles at two loci are preferentially associated in gametes (LD), the random association of gametes yields an excess of either double-homozygous or double-heterozygous genotypes (ID).

is misleading: AO can emerge without overdominance, although it does require a statistical association between heterozygosity at marker loci and heterozygosity at fitness loci in diploid individuals. This statistical association is called identity disequilibrium, which should not be confounded with linkage disequilibrium, which occurs in the gametes (see Table 1).

THE ORIGIN OF HFCs: THEORY AND PRACTICE

HFCs do not arise in large random-mating populations at equilibrium

HFCs are widespread in natural populations of a variety of organisms (Mitton and Grant 1984; Zouros and Foltz 1987). Available data show, however, that HFC is typically a weak signal—the explained variance in fitness traits is usually a few percent or less (Britten 1996; David 1998; Coltman and Slate 2003; Chapman et al. 2009). Significant correlations do emerge from time to time, but only when sample size and population contexts are favorable, rather than in a consistent manner within a species. Thus, although polymorphisms causing inbreeding depression in controlled crosses abound in most genomes, it is only in particular contexts that they will generate detectable HFC in natural populations.

A recurrent debate in the HFC literature focuses on the extent to which marker heterozygosity reflects a more general state of the genome, that is, genome-wide heterozygosity. In a large (e.g., $N \sim 500$), random-mating population at equilibrium, it is clearly not the case, as theoretically derived by Ohta (1971) and Ohta and Kimura (1970), restated with empirical examples by Houle (1989) and Whitlock (1993), and simulated by Balloux et al. (2004). Indeed, in the absence of linkage and identity disequilibria (LD and ID, respectively, Table 1), all loci within genomes come to be heterozygous or homozygous independently of each other. Thus, the only source of variance in the proportion of heterozygous loci in genomes is pure binomial sampling variance, which is negligible as the number of polymorphic loci is high (Chakraborty 1981). With no source of interlocus correlation, MLH at a set of markers will only reflect the state of these particular markers, and will not be correlated to heterozygosity elsewhere in the genome (Chakraborty 1981, 1987; Lynch and Walsh 1998; Slate et al. 2004; Fig. 1).

For marker loci to indicate anything other than their own state, and thus for HFC to arise, one of the three following processes (or any combination thereof) must occur: (1) a fraction of systematic consanguineous matings (Ohta and Cockerham 1974), (2) genetic drift, of concern to small populations (Ohta 1971), or populations recently affected by bottlenecks (Bierne et al. 2000a); in this case more or less consanguineous matings occur at random because in a finite population, the degree of relatedness varies among random pairs of individuals, and (3) admixture or immigration (Tsitrone et al. 2001); in this case individuals with mixed ancestry (i.e., ancestors from different populations) are relatively outbred compared to "pure" genomes (relatively inbred). All three mechanisms represent some form of variation in inbreeding sensu lato, and generate ID. LD are not required to generate HFC and are not produced by systematic consanguineous matings, because the latter only modify the way gametes are united to each other, and do not affect their frequencies (Charlesworth et al. 1991; Vitalis and Couvet 2001). With drift or immigration, LD is produced (Hill and Robertson 1968; Sved 1968; Bierne et al. 2000a) and ID emerges as a consequence of LD (occurring in the gametes) and random mating. Note that although they do not produce LD by themselves, nonrandom matings can reinforce LD (and ID) produced by drift or migration because they decrease effective recombination rates.

The genesis of HFCs, under any of the three scenarios, can be formalized in a very simple way, detailed below. This model

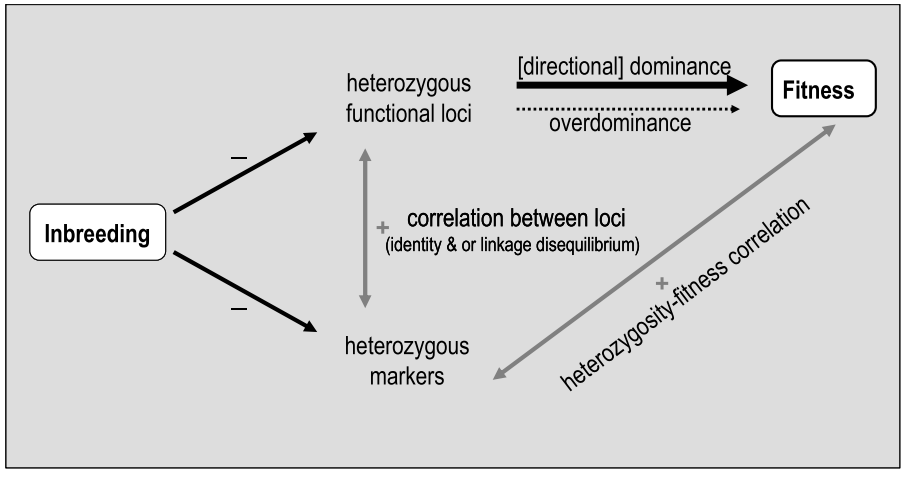

Figure 1. **A path diagram showing how HFCs arise from inbreeding. Black arrows represent a causal relationship, gray arrows a statistical correlation.**

assumes that marker loci are neutral (e.g., noncoding DNA markers) and that inbreeding depression exists, although it makes no assumption about its basis (directional dominance vs. overdominance). We will first assume no physical linkage among loci; this is not absurd because most pairs of loci in a genome are either not or only very loosely linked. We will come back later to the effect of physical linkage (Section "Variation in HFC among Loci and the Dichotomy of Local versus General Effects"). We will also assume that all the variance in genomic heterozygosity among individuals can be represented by a single variable f, which is determined for each individual by its pedigree, and represents the probability for any locus to be autozygous (i.e., homozygous and identical by descent) in that individual.

Correlations in heterozygosity among loci and how to measure them

We have seen that ID is the fundamental cause of HFC. ID can be estimated, and its significance can be tested, using a sample of multilocus genotypes. A popular procedure to test the significance of ID is to divide a set of loci into two halves and to compute the correlation between MLH in the first half and MLH in the second half (heterozygosity–heterozygosity correlations or HHC, Balloux et al. 2004); this can be repeated using different partitions of the set of loci. However, this procedure yields a complicated distribution of HHC coefficients, which are not independent from one another, and provide no synthetic measure that can be related to HFC theory. The central measure of ID in HFC theory is the excess of double heterozygotes at two loci relative to the expectation under random association (i.e., covariance in heterozygosity), standardized by average heterozygosity. Under any form of inbreeding, this measure is constant whatever the pair of loci (i, j) considered, and equals a parameter named g_2, which depends only on the mean and variance of inbreeding in the population, and not on locus-specific characteristics (David et al. 2007):

$$\forall(i,j),\ \frac{\mathrm{cov}(h_i,h_j)}{\bar{h}_i\bar{h}_j}=\frac{E(h_ih_j)-\bar{h}_i\bar{h}_j}{\bar{h}_i\bar{h}_j}=g_2=\frac{\sigma^2(f)}{(1-\bar{f})^2}, \quad (1)$$

where h_i is the heterozygosity at locus i($h_i = 1$ for an heterozygote, 0 for an homozygote), $E(h_ih_j)$ is the mean of h_ih_j, and $\sigma^2(f)$ and $\bar{f}$ are the variance and mean of f among individuals.

All loci can be combined to produce a single estimate of g_2 (e.g., Slate et al. 2004). The procedure is detailed in David et al. (2007), who provide a freeware to compute this estimate, and to test whether g_2 differs significantly from zero (the software allows for missing data; it also computes selfing rates under the hypothesis that ID is due to partial selfing—but g_2 estimates are valid with any source of inbreeding). g_2 can be represented graphically as an excess of multiheterozygous and multihomozygous genotypes, inflating the variance in MLH, compared to a random assortment of loci (Fig. S1). Without variance in inbreeding (i.e., when $g_2 = 0$), HFC cannot arise. However, nonsignificant values of g_2 do not contradict the observation of HFC. Indeed, fitness traits capture the effect of potentially many more loci than the number of markers typed. Therefore a slight inbreeding is often more easily detected through its effect on the phenotype than through its effect on heterozygosity at a few marker loci. Thus, HFCs reach the significance threshold more easily than correlations in heterozygosity in pairs of loci (g_2). However, g_2 remains a more powerful statistic than HHC to detect ID, as it uses information on associations from all loci simultaneously. Moreover, g_2 is also a central mathematical entity in HFC theory (see below).

HFC is the product of two correlations

Assuming neutral markers, correlations between heterozygosity (h) and fitness (W) arise from simultaneous effects of inbreeding level f on these two variables (Fig. 1)

$$r(W,h)=r(W,f)r(f,h). \quad (2)$$

A similar argument holds for regression slopes of fitness on heterozygosity

$$\beta(W,h)=\beta(W,f)\,\beta(f,h). \quad (3)$$

The correlation between marker heterozygosity and inbreeding

We now detail the first of the two right-hand terms of equations (2) and (3). The correlation between h and f can be computed with a simple partition method. The population is partitioned into inbreeding classes within which no covariance in heterozygosity exists among loci. One then obtains (Bierne et al. 2000a; Slate et al. 2004)

$$\mathrm{cov}(h_i,f)=-h_{0,i}\sigma^2(f), \quad (4)$$

where $h_{0,i}$ represents the expected heterozygosity at locus i in the absence of inbreeding (gene diversity in the reference generation). From this expression, expected regressions and correlations between h_i (or multilocus heterozygosity H, the sum of h_i's over a set of loci, or any standardized measure of MLH) and f can be easily derived—they are presented in Table 2. All these quantities are negative and increase in intensity with increasing g_2. These theoretical expectations are highly concordant with observed correlations between heterozygosity and f when the latter can be estimated independently using pedigree data (Fig. 2). The regression slope of observed on expected correlations is close to unity, although both heterozygosity and pedigree f are estimated with sampling error, resulting in scatter around the regression line.

Table 2. Expected components of the relationship between various measures of marker heterozygosity (x) at L loci and the inbreeding level under a simple inbreeding model without linkage. The various measures include heterozygosity at a single locus i (h_i, coded as 0 or 1), multiple locus heterozygosity (H, the sum of h_i over L loci), and various standardized measures often used by authors: the relative heterozygosity at a single locus ($h_i^*=h_i$ divided by its average value in the population), relative heterozygosity at multiple loci (H^*=sum of h_i^* over L loci), and standardized multilocus heterozygosity ($H^{**}=H$ divided by its population mean). Note that the H in Slate et al. (2004) is equivalent to our H^{**}. Formulae are given for the expected covariance cov(f, x) between inbreeding f and each of the five measures (x), their expected correlation coefficient ($r(f, x)$), the slope of the regression that predicts f as a function of x ($\beta(f, x)$), and the variance in x. The latter is given for consistency in the formulae but can usually be directly estimated from data. Mean and variance operators for a given variable y are noted $\bar{y}$ and $\sigma^2(y)$, respectively. Most expressions are given as functions of g_2, defined in equation 1.

$x=$	cov(f, x)	$r(f, x)$	$\beta(f, x)$	$\sigma^2(x)$
h_i	$-h_{0,i}\sigma^2(f)=-\bar{h}_i g_2(1-\bar{f})$	$-\frac{\bar{h}_i}{\sigma(h_i)}\sqrt{g_2}$	$-\frac{\bar{h}_i}{\sigma^2(h_i)}g_2(1-\bar{f})$	$\bar{h}_i(1-\bar{h}_i)$
$H=\sum_i h_i$	$-\sum_i h_{0,i}\sigma^2(f)=-\bar{H}g_2(1-\bar{f})$	$-\frac{\bar{H}}{\sigma(H)}\sqrt{g_2}$	$-\frac{\bar{H}}{\sigma^2(H)}g_2(1-\bar{f})$	$\sum_i \sigma^2(h_i)+\left(2\sum\sum_{j>i}\bar{h}_i\bar{h}_j\right)g_2$
$h_i^*=h_i/\bar{h}_i$	$-\frac{\sigma^2(f)}{(1-\bar{f})}=-g_2(1-\bar{f})$	$-\frac{\bar{h}_i}{\sigma(h_i)}\sqrt{g_2}$	$-\frac{g_2(1-\bar{f})}{\sigma^2(h_i)}$	$\frac{1-\bar{h}_i}{\bar{h}_i}$
$H^*=\sum_i h_i^*$	$-L\frac{\sigma^2(f)}{(1-\bar{f})}=-Lg_2(1-\bar{f})$	$-\frac{L}{\sigma(H^*)}\sqrt{g_2}$	$-\frac{L}{\sigma^2(H^*)}g_2(1-\bar{f})$	$\sum_i \sigma^2(h_i^*)+L(L-1)g_2$
$H^{**}=\frac{H}{\bar{H}}$	$-\frac{\sigma^2(f)}{(1-\bar{f})}=-g_2(1-\bar{f})$	$-\frac{\sqrt{g_2}}{\sigma(H^{**})}$	$-\frac{g_2(1-\bar{f})}{\sigma^2(H^{**})}$	$\frac{\sigma^2(H)}{\bar{H}^2}$

The correlation between inbreeding and fitness

The correlation between inbreeding and fitness (second right-hand term of eq. 2) depends on three variables: (1) the slope of the regression line ($\beta(W,f)$, or inbreeding load), which represents the effect of inbreeding on fitness, and depends on the average number of deleterious recessives per diploid genome, (2) the variance in inbreeding in the population $\sigma^2(f)$, and (3) the variance in fitness due to environmental or genetic effects that are not related to inbreeding (e.g., additive genetic variance) in the population. Thus,

$$r(W, f) = \beta(W, f)\frac{\sigma(f)}{\sigma(W)}. \quad (5)$$

Empirical data on heterozygosity and fitness, together with equations (2)–(5) and Table 2, are sufficient to estimate both the inbreeding load $\beta(W, f)$ and the amount of variance in fitness explained by inbreeding $r^2(W, f)$ which are the quantities of interest to researchers studying the impacts of inbreeding in natural populations, We recommend that they should be routinely computed in future HFC studies. Worked examples using cassava and shrimp datasets from Pujol et al. (2005) and Bierne et al. (2000b) are provided in the Appendix S1. Note that HFC itself ($r(W, h)$) is not directly informative on the actual impact of inbreeding as it only reflects how well real data are approximated by a regression of fitness on heterozygosity. It does not quantify the strength of inbreeding depression ($\beta(W, f)$).

WHAT META-ANALYSES TELL US ABOUT HFC

The above theory predicts that even when a large fraction of the variance in fitness is due to inbreeding, $r^2(h, f)$, and therefore $r^2(W, h)$, will be very small under realistic levels of inbreeding in natural populations unable to self-fertilize (as in the data from Fig. 2). One may wonder how often significant HFC can arise in true datasets. Several meta-analyses have summarized existing information on the effect size of HFCs (Britten 1996; Coltman and Slate 2003; Chapman et al. 2009). All of them show that HFC is usually very weak. The overall percentage of variance in fitness that MLH explains ($r^2(W, h)$) spans from 0.07% to 3.3%, with an average of 1% or less for microsatellite studies; Chapman et al. 2009). This is consistent with inbreeding-based theory; indeed, both particular population structures (large variance in inbreeding coefficients) and high statistical power are needed to obtain significant HFCs. Moreover, even under such circumstances, HFC represents only the tip of the inbreeding iceberg (i.e., $r^2(W, h) \ll r^2(W, f)$ because $r^2(f, h) \ll 1$, eq. 2).

Several elements can influence the estimates of effect sizes of HFCs. The first is publication bias (Coltman and Slate 2003; Chapman et al. 2009). This problem is mitigated by the frequent publication of many nonsignificant correlations, alongside the rarer significant ones, for each studied system. Moreover, some studies of HFC may have been undertaken as an "afterthought," that is, microsatellites have been typed for other purposes (e.g., paternity assignments, evaluation of genetic structure) and later related to fitness traits. In these cases, the studied populations may

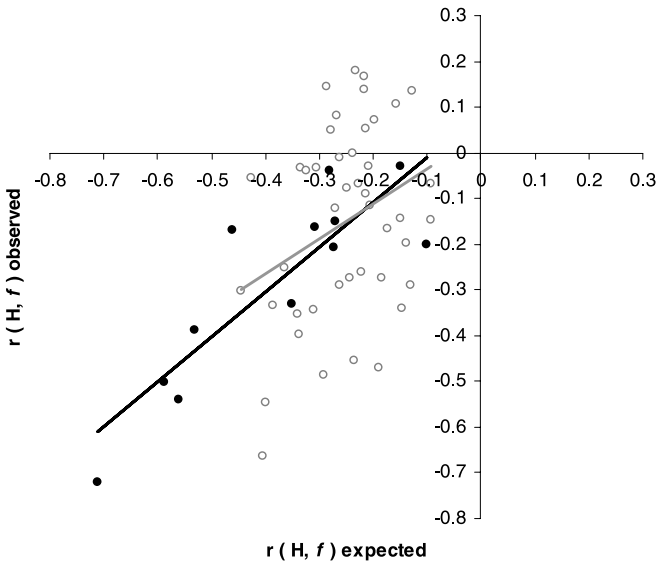

Figure 2. Observed and expected values of the correlation between standardized heterozygosity and inbreeding *f* . Black symbols and regression line refer to vertebrate studies ($n = 12$, mean/median sample size per study = 181/120). Populations sampled: Coopworth sheep (*Ovis aries*), Bighorn sheep (*O. Canadensis*), Wolves (*Canis lupus*, captive), Lipizzan horses (*Equus caballus*), Large ground finch (*Geospiza magrinostris*), Medium ground finch (*G. fortis*), Cactus finch (*G. scandens*) based on Slate et al. (2004), Wolf (*Canis lupus*, wild, Bensch et al. 2006), Siberian jays (*Perisoreus infaustus*, Alho et al. 2009), Soay sheep (*O. aries*, Overall et al. 2005), Icelandic sheepdog (*C. familiaris*, Olafsdottir and Kristjansson 2008), Gazelle (*Gazella dorcas neglecta*, Ruiz-Lopez et al. 2009). Open gray symbols and regression line represent dog breeds presented in Leroy et al. (2009) ($n = 44$ dog breeds, mean/median sample size per breed = 24/24)—note that small sample sizes for individual datapoints will generate greater stochasticity in observed $r(H, f)$. The slope of the regression line for $r(H, f)$ expected and observed is $\beta = 0.983$ (+0.181 SE) for vertebrate studies and $\beta = 0.771$ (+0.347 SE) for dog breeds. Both slopes differ significantly from zero ($F_{1,10} = 29.49$, $P < 0.001$, and $F_{1,42} = 4.94$, $P = 0.032$, respectively), yet as expected, they do not differ significantly from each other ($T_{11} = 1.15$, $P = 0.274$).

not have any of the genetic structures required to generate HFCs (Section "What Meta-Analyses Tell Us about HFC"). The second source of variation in HFC estimates lies with the choice of the phenotypic trait. In principle, only traits that exhibit directional dominance, and are affected by many loci, can correlate with heterozygosity. Integrated life-history traits (e.g., survival, reproductive success), and growth traits (growth rate or size at a juvenile stage) involve many different loci, all of which are targets for deleterious recessive mutations (Houle et al. 1996; Houle 1998). This genetic architecture favors the expression of HFC. Morphological (e.g., organ size in adults, meristic traits) and behavioral traits may often be under stabilizing rather than directional selection, they may display less-directional dominance, and be affected by fewer loci than life-history traits (Houle et al. 1996); therefore they should rarely be correlated to heterozygosity. In practice, HFC is often tested using all possible traits, thereby reducing the term HFC to a misnomer (Chapman et al. 2009).

VARIATION IN HFC AMONG LOCI AND THE DICHOTOMY OF LOCAL VERSUS GENERAL EFFECTS

Theory predicts, and meta-analytical studies confirm, that HFCs are weak. This makes HFC studies prone to "fishing trips," where the same data are used in different analyses until some significant effect is found. In particular, it is tempting to analyze each locus separately to find correlations that stand out. Such estimates of HFC at individual loci have been considered as indicative of local effects (Hansson and Westerberg 2002; Hansson et al. 2004; Da Silva et al. 2006; Lieutenant-Gosselin and Bernatchez 2006; Charpentier et al. 2008; Hansson and Westerberg 2008; Vilhunen et al. 2008; Da Silva et al. 2009). However, speculations on local effects have generally not been supported by relevant and significant statistics. Single-locus HFCs need to be interpreted cautiously, as discussed below. A statistical method allowing parsimonious identification of local effects is also provided (subsection "Statistical Issues: Appropriate Testing for Local Effects").

Local and general effects: definition and historical background

The terms "local" and "general" effects, originally coined by David et al. (1995), are deeply rooted in the historical context of the 70s to the 90s (Kimura 1983; Ohta 1992), when allozymes were the only available markers. At that time the debates on HFC opposed two hypotheses: direct overdominance (a direct effect of allozyme loci on the phenotype) and AO. In this context, the term "local effects" sought to group together: direct effects of marker genes and indirect effects of loci in the chromosomal vicinity of the markers, by contrast with "general effects," to which the entire genome might contribute. This grouping was made purely for practical reasons, as it was impossible to disentangle whether the phenotypically active loci were the markers themselves or closely linked loci. Both terms survived the transition to microsatellite markers, which are—but for a few exceptions—noncoding polymorphic loci (Jarne and Lagoda 1996). In this case, local effects can only reflect the action of coding loci physically linked to a particular microsatellite (as direct effects are excluded). In the following sections, we will not discuss direct effects (irrelevant to microsatellites) but will concentrate on local effects of the "indirect" type, involving physical linkage.

Local effects are unlikely to be detected

Local effects imply that most or all of the HFC observed at a particular marker locus is due to polymorphic fitness genes in the chromosomal vicinity of that marker. This requires ID to be

on average higher between two linked loci than between two unlinked loci, so that HFC at the marker locus will capture a higher proportion of the phenotypic variance generated by a linked locus compared to an unlinked locus. This property is predicted by most inbreeding models of HFC (Ohta 1971; Bierne et al. 2000a). However the number of loci linked to a particular marker is usually much smaller than the number of unlinked loci; local effects will be important only if the difference between linked and unlinked loci is strong enough to compensate this asymmetry. Otherwise, the cumulative effects of fitness loci in all unlinked parts of the genome may override that of the chromosomal region of the marker.

Theory predicts that this dilution effect will occur in all the contexts that can generate HFC, so that local effects will usually be very difficult to detect. In the case of recent inbreeding (e.g., partial selfing, or mating between relatives), ID (hence HFC) is expected to be at most twice as large for two completely linked loci as for unlinked loci (from Weir and Cockerham 1973, eq. 37). This ratio of less than 2 is not sufficient to counter-balance the fact that linked loci typically represent no more than a few percent of the genome. However, higher ratios can be found under other demographic scenarios, such as small population sizes, bottlenecks, or admixture. In these cases, ID is associated with LD, which decreases with physical distance (Hill and Robertson 1968). A favorable situation to observe local effects would therefore arise if LD were absent among unlinked loci in a population while being strong among linked loci. Unfortunately, such situations are exceedingly rare: population bottlenecks and population admixture do generate moderate levels of linkage disequilibrium between physically unlinked loci, although they are lower than for linked pairs. In real datasets, the proportion of variance in LD among markers that can be explained by physical distance typically ranges from a few to 45% (Slate and Pemberton 2007). Yet unlinked markers (nonsyntenic, placed on different chromosomes) frequently exhibit LD with each other (e.g., Slate and Pemberton 2007; Li and Merilä 2009).

To model local effects in HFC, not only the marker loci but also fitness loci must be included in simulations. Unfortunately, models investigating the potential role of local effects have sometimes limited their simulations to neutral markers, potentially generating biased conclusions (Balloux et al. 2004; Hansson and Westerberg 2008). Theoretical approximations and simulations including both marker and fitness loci show that (1) AO can arise even with no linkage at all, as stated by Ohta in the 1970s (Ohta 1973; Ohta and Cockerham 1974) and (2) with realistic genome sizes and conditions, the contribution of physically linked loci is on average small when compared to the cumulative effect of the entire genome (Ohta 1973; Bierne et al. 2000a). Based on earlier work (Bierne et al. 2000a), Figure 3 presents simulations assessing the relative contributions of linked and unlinked detrimental alleles in a genome of standard vertebrate recombinational size. Clearly, detrimental alleles that are physically linked to a marker produce stronger AO (approximately 2.5 times) than the same number of unlinked detrimentals (Fig. 3A), consistently with previous reports (Ohta 1973; Ohta and Cockerham 1974; Bierne et al. 2000a). However, in a genome of 20 chromosomes, 95% of the genome is unlinked to a particular marker. In consequence, the increase in HFC due to linkage is diluted and barely detectable (Fig. 3B).

At first thought, an appealing scenario to generate local effects is to consider a large panmictic population in which LD would be generated by selection, at a very small chromosomal scale. This could occur if the marker happened to fall in the vicinity of either an overdominant locus or of a locus with ongoing fixation of a favorable allele (i.e., a case of genetic hitchhiking; Maynard-Smith and Haigh 1974; Slatkin 1995; Schierup et al. 2000). Such a scenario is, however, extremely unlikely to contribute to any observed HFC because the polymorphism under selection would have to be very recent, very strongly selected, and very close to the marker locus. Although selection can leave a footprint on an appreciable proportion of some genomes (Nielsen et al. 2005; Li and Stephan 2006), adaptive evolution probably occurs only at a few segregating mutations at any given time (Kreitman 2000; Hedrick 2006b;). In *Drosophila* for instance, it has been inferred that one advantageous mutation reaches fixation every ~500 generations (Eyre-Walker 2006) and that the average fitness effect is no more than 0.5% (Li and Stephan 2006). Thus, only a few favorable mutations are on their way to fixation at any one time and they affect only a small chromosome fragment (~no more than 100 Kb, Li and Stephan 2006), making them very unlikely to be detected in an HFC context. Balanced polymorphisms are also known not to thrive in genomes and they are often too old to retain linkage disequilibrium on a large chromosomal scale (Charlesworth 2006). In any case, the fitness effects of all these adaptive polymorphisms, detected by population genetic methods (integrating the effect of selection on many generations), are far too small (typically $s < 1\%$, Li and Stephan 2006) to be detected individually through a genotype/fitness correlation in a single generation. On the contrary, the cumulative effect of many detrimental alleles distributed along the entire genome can produce fitness differences that are strong enough to generate detectable HFCs.

In summary, in realistic situations the effect of linked mutations on HFC will be overridden by the rest of the genome in natural populations. However, the effect of the linked genetic load around a marker can be revealed using special crossing designs or analyses. In such cases, variation in genome-wide heterozygosity is reduced as much as possible to compare individuals with the same genomic background on average—ideally individuals with identical pedigrees (i.e., full-siblings) (Charlesworth

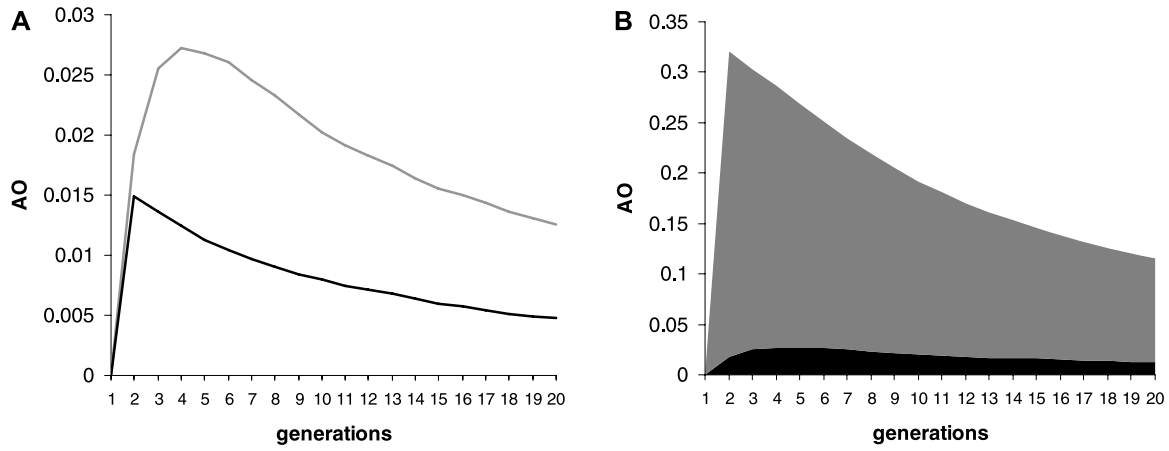

Figure 3. **The signal of local effects in a vertebrate population undergoing a demographic bottleneck is easily lost due to cumulated effects of detrimental alleles scattered across the genome (general effects). Associative overdominance (AO, fitness advantage of heterozygotes relative to homozygotes) was simulated in a standard vertebrate genome of 20 chromosomes and a map length of 1 M (a total genome size of 20 M), based on a model by Bierne et al. (2000b). Deleterious mutations were randomly dispatched along the genome with U (genomic deleterious mutation rate) = 2, h (dominance coefficient) = 0.1, s (selection coefficient) = 0.05. In a large population at equilibrium, the number of deleterious mutations per genome is thus: $n = U/(h \times s) = 400$. We modeled a large, monogamous vertebrate population that underwent a sustained population bottleneck. Forty individuals taken from this population (Generation 0) found a small population that remains at $N = 40$ for all subsequent generations; mating between relatives can occur in generation 1, thus producing inbred individuals for the first time in generation 2. The AO at one marker locus was investigated for 20 generations. To decompose the effect of linked and unlinked deleterious mutations, three scenarios were used: (1) local effects—the marker is localized within a single chromosome of 1 M with 1/20th of the load ($U = 0.1$), (2) general effects using one chromosome only: the marker segregates independently (it is localized in a chromosome freed from deleterious mutations) in a genome of 1 M with 1/20th of the load, and (3) general effects using the full genome size: a marker segregates independently in a genome of 19 M with 19/20th of the load ($U = 1.9$). (A) AO generated by one chromosome when the marker is on the same chromosome (gray line) and when it is on a different chromosome with no deleterious mutations (black line). (B) Total AO partitioned into AO caused by detrimentals located on the same chromosome as the marker (black area) and caused by detrimental alleles occurring on the remaining 19 chromosomes (gray area).**

and Charlesworth 1999). Importantly, these individuals need to be inbred to some degree to expose the phenotypic effects of recessive mutations. Analyses of large sibling groups from crosses between related parents, with or without molecular markers, are classical methods to explore the genetic architecture of inbreeding depression (Hedrick and Muona 1990; Fu and Ritland 1994; Bierne et al. 1998; Carr and Dudash 2003; Haag and Ebert 2007). Natural pedigrees, in combination with molecular markers, can be used in the same way. This has been done by comparing sibling dyads from several families of reed-warblers (Hansson et al. 2001, 2004). In these studies, the heterozygosity in surviving offspring was greater than in their nonsurviving siblings. The background variation at unlinked loci (genome-wide heterozygosity) was eliminated because both siblings have the same inbreeding coefficient. Therefore, HFC was restricted to fitness genes physically linked to a marker locus for each particular dyad. However, (1) some inbreeding sensu lato must be present in some of the dyads, otherwise fitness would be correlated to particular alleles rather than heterozygosity (2) not all sibling dyads are equally inbred, and they do not necessarily carry mutations in the vicinity of the same markers. Thus, what HFC represents in this case is a "sample" of the cumulative effects of the linked load present in some families that happen to be slightly inbred. Although these studies can determine quite powerfully whether a more heterozygous sibling is more likely to survive, chances of pin-pointing fitness genes through local effects, given that each dyad will carry different ID and LD signatures, are limited.

We have argued that, without a specific design (i.e., full-siblings dyads), local effects are expected to be small and very hard to detect. Yet, in many cases, only some loci within a given set of markers show a significant single-locus HFC, a variation that is often interpreted as a consequence of local effects. We believe that this apparent contradiction can be resolved using appropriate statistical tests.

Statistical issues: appropriate testing for local effects

As stated earlier, the problem is not whether local effects exist (all models predict higher contribution of linked loci to HFC) but whether a large, or even detectable, part of the overall correlation is due to local effects. It is relatively easy and tempting to test, locus by locus, whether the slope of the regression of fitness on single-locus heterozygosity ($\beta(W, h_i)$) differs significantly from

zero. However, such approach does not test for local effects. Under the general effect hypothesis, all $\beta(W, h_i)$'s are predicted to differ from zero. Because of sampling variance, slope estimates for particular loci will differ from each other. Given that the correlations are usually very weak and therefore estimated with substantial error, both positive and negative estimates are to be found, and some of them may significantly differ from zero whereas others may not. On average (across different loci) single-locus regression estimates are expected to be weakly positive. The crucial point to address is not whether these slopes differ significantly from zero, but whether some regression coefficients differ from other single-locus regression coefficients. To implement this test, one cannot use separate regressions for each locus as they are not independent from each other, owing to ID (Table 4). The appropriate procedure is to test whether a multiple regression incorporating specific effects for each locus explains more variance than a simple regression on MLH (David 1997). This can be achieved by performing the following steps: (1) Regress fitness trait (W) on MLH (H in table 2), using a simple regression: Model m_1: $\mathrm{W} = a\,H + b$, with $p_1 = 2$ estimated parameters; (2) Regress fitness on all single-locus heterozygosities $h_i \ldots h_{L,}$ expressed as one or zero (L being the number of loci), using a multiple regression: Model m_2: $\mathrm{W} = a_1h_1 + a_2h_2 + \ldots + a_Lh_\mathrm{L} + b$, with $p_2 = L + 1$ parameters; and (3) Test whether the two models differ significantly from each other using an F-ratio test. Without local effects, we expect all marker loci to have the same partial regression coefficients in the multiple regression of the fitness trait (H_o: $a_1 = a_2 = \ldots a_L$). By contrast, if local effects occur, one or more partial regression coefficients (corresponding to markers located near one or several genes with large fitness effects) are expected to differ from the others. H_o should then be rejected.

Significant local effects can be identified if m_2 explains significantly more variance than m_1, that is, if the residual sum of squares (SS) of m_2 is significantly larger than the residual SS of m_1:

$$F = \frac{(resSS_1 - resSS_2)/(df_1 - df_2)}{resSS_2/df_2},$$

where $df_1 = N$(sample size) $- 2$ and $df_2 = N - L - 1$ are the residual degrees of freedom for m_1 and m_2, respectively. The F-ratio should be compared with tabulated F values with $(df_1 - df_2)$, df_2 degrees of freedom—a value greater than the tabulated value allows for the rejection of the null hypothesis. The same test can be applied to nonnormally distributed fitness variables using generalized linear models (e.g., in the case of fitness variables such as survival, fecundity, infection status etc)—the residual SS are simply replaced by residual deviances. Moreover, this method should only be applied when there are no major differences in observed heterozygosity among loci; otherwise the data should be normalized as explained in the Appendix S2. Note that the power of this test (and the accuracy of the multiple regression) decreases when the number of loci approaches the number of individuals.

Another important prerequisite for the test is that the genotypes need to be complete, and exactly the same set of individuals and loci must be included in the two models (otherwise the SS for the different tests will differ). As incomplete genotypes are frequent, a simple remedy is to replace missing data (i.e., loci for which the heterozygosity h_i is unknown for a particular individual in the sample) by the sample average ($h_i = \bar{h}_i$), in both models (Table S1). These "filled" missing values are conservative to the partial regression coefficient at that locus (as they do not bring any information on the status of the individual relative to the population average), but they do allow to make use of the information contained at other loci (and to increase the overall power of the test) rather than discarding many individuals from the analysis.

Although the above procedure allows for rigorous testing of local effects, it is undeniable that the latter are extremely difficult to detect if the MLH regression (m_1) is already weak. Indeed, we are not aware of any HFC data that passed this test and detected significant local effects.

HFCs AND ALTERNATIVE MEASURES OF INBREEDING: FUTURE DIRECTIONS

Although their theoretical underpinnings are simple, HFCs are not straightforward to measure in practice. A summary of common misconceptions about HFCs, and a practical guide for analysis and interpretation are presented in Tables 3 and 4. Below, we outline how HFCs, if correctly interpreted, may provide insights into the impacts of inbreeding depression in populations, and discuss the complementarity of HFCs and other approaches in this context.

As illustrated in the examples of Appendix S1, HFCs are a convenient, if rather imprecise, way to quantify the impact of inbreeding on fitness in natural populations.

Ideally, combining molecular markers with controlled crosses is a more powerful approach to quantify the genetic load (Bierne et al. 1998; Launey and Hedgecock 2001), as well as to localize and characterize polymorphisms responsible for inbreeding depression (e.g., Fu and Ritland 1994; Haag and Ebert 2007). Yet, controlled crosses are not always feasible (e.g., in long-lived vertebrates), or may be restricted to a few families (Remington and O'Malley 2000). As deleterious alleles are rare and scattered in genomes (Charlesworth and Charlesworth 1987), additional families would harbor different assemblies of deleterious alleles. In addition, controlled crosses yield estimates of the potential rather than the actual impact of inbreeding on fitness in natural populations, precisely because the degree of inbreeding is artificially imposed. In addition to being the only available option in

Table 3. A summary of common misconceptions about HFCs.

Assertion	Reality
Heterozygosity does not estimate inbreeding in real populations.	Yes it does—but with decreasing precision when inbreeding decreases. A correlation between individual heterozygosity and inbreeding appears as soon as inbreeding variance is generated owing to nonrandom mating, immigration or drift, although the correlation may be low and transient.
As HFCs cannot be explained by inbreeding, linkage to genes under balancing selection explains HFCs.	Physical linkage is not an alternative to inbreeding. Indeed, linkage enhances but does not generate associations between loci (Hartl and Clark 2007). HFCs require a correlation in heterozygosity between marker and fitness loci that inevitably implies some sort of inbreeding. Correlations in heterozygosity will affect the entire genome, although they will usually be stronger between linked loci. In addition, the nature of selection (balancing or not) cannot be assessed from HFC; Moreover deleterious mutations are arguably more frequent than balanced polymorphisms within genomes.
The opposition between "local" and "general" effects was introduced to contrast inbreeding and linkage as the source of HFCs.	These terms were introduced at a time when direct selection on allozymes was a serious hypothesis to explain HFCs as noncoding DNA markers were not yet available (David et al. 1995). Statistical procedures were developed at the time to test the hypothesis of direct selection; however, because no test could distinguish between direct effects of allozymes and potential indirect effects of closely linked loci, the "local" versus "general" alternative was substituted to "direct" versus "indirect" (David 1997).
ID and LD reflect two different processes, one operating in large populations with nonrandom mating and the other in finite populations at a small chromosomal scale.	No. ID and LD are two measures of departure from random associations between loci. ID and LD are often jointly generated. Only consanguineous matings in large populations can generate ID without LD. In small, bottlenecked or admixed population, LD co-occurs with ID because if alleles at two loci are preferentially associated in gametes (LD), the random association of two gametes will yield an excess of either double homozygous or double heterozygous genotypes (ID) (Yang 2000). In this case LD and ID can be considered as two faces of the same coin. HFC in small populations can be equivalently modeled using LD (Ohta 1971) or ID (Bierne et al. 2000b).
Only closely linked fitness loci can have an effect on a given locus.	It is true that closely linked loci have a stronger effect on HFC at a given marker than unlinked loci. However, the sum of many unlinked loci, each with a small effect, is expected to exceed the contributions of the few linked loci (see Section "Variation in HFC among Loci and the Dichotomy of Local versus General Effects").
HFC generated by LD can only be interpreted as local effects.	No. In small or nonequilibrium populations, LD can also be generated between physically unlinked loci (e.g., Slate and Pemberton 2007).
HFC can be used as an alternative to LD mapping to find genes that contribute to a given phenotype.	No. LD mapping is different from HFCs. In LD mapping, a population sample is split in groups based on phenotypic features, and differences in allelic frequencies between groups are looked for. HFC reveals the action of genes that cause inbreeding depression on the phenotype within a population; yet the usually very sparse coverage of the genome by markers, the fact that inbreeding is often low within populations, the multiplicity of loci that contribute to inbreeding depression, and the relatively small effect of physical linkage on HFC make it unrealistic to map quantitative trait loci using HFC.

some circumstances, HFC has the advantage of yielding overall estimates of both the potential genetic load and its actual impact in a population, its main disadvantage being its low precision and the subsequent need for very large samples.

Pedigree analysis provides direct estimates of individual inbreeding coefficients, and when feasible is an attractive alternative to HFC. Although both approaches have their strengths and limitations, particular contexts may favor one over the other. In principle, accurate and deep pedigrees offer greater precision in estimating individual inbreeding coefficients and hence inbreeding depression (Balloux et al. 2004; Pemberton 2004; Slate et al. 2004). However, this accuracy heavily relies on the validity of the assumption that founding individuals (the end-points of the pedigree) are outbred and unrelated. Often this assumption is violated, and relatedness/inbreeding estimates are biased downwards (Ruiz-Lopez et al. 2009). To minimize bias, it is important to restrict datasets so that individuals identified as inbred or outbred have enough known ancestry (e.g., at least two generations in species with separate sexes) to justify their classification. This can drastically reduce sample sizes (Marshall et al. 2002; Szulkin et al. 2007). Finally, pedigree is not obtainable in highly mobile or fecund species, and in species with large, open, or inaccessible

Table 4. A practical guide on how to analyze HFCs.

Step	What to do
1. What traits to use?	Focus on traits that are the product of a large number of loci, and/or that are closely related to fitness (Section "What Meta-Analyses Tell Us about HFC"). Traits may be scaled in various ways depending on their distribution, and on the assumed additive or multiplicative nature of mutational effects on the phenotype. In many cases, the model will not be a simple linear regression but a generalized linear model, in which case appropriate error treatment is required (i.e., survival fitted as binomial/quasi-binomial variable with logit link, fecundity as Poisson/quasi-Poisson with log link, growth rate and body/organ size with identity link and normal errors).
2. Estimating heterozygosity from molecular data	Several measures of heterozygosity can be used, generating equivalent results: H (a simple count of the number of heterozygous loci) is the simplest, but various standardizations are available (see Table 2). Always state clearly which standardization you use! While other measures of inbreeding such as internal relatedness (IR), adaptive distances (Smouse 1986; Houle 1994), or γ (David 1997) may be theoretically more precise, as they take into account allelic frequencies, scoring errors and especially null alleles may generate substantial outliers because apparent homozygotes for rare alleles have very extreme values, and have a relatively high probability to be heterozygous in reality. As datasets are never error free, MLH is in general the most robust measure (see also Chapman et al. (2009) for discussion). The use of d^2 with microsatellites, which takes into account size differences between alleles in a heterozygote, is not justified because it is not as good a predictor of inbreeding as H (Tsitrone et al. 2001).
3. Estimating the strength of ID in the sampled population	Test for ID using g_2 (Section "Correlations in Heterozygosity among Loci and How to Measure Them", Appendix 1, and David et al. 2007). This procedure not only tests, but also provides a measure of genetic associations that can be related to HFC theory; it handles missing data.
4. Assessing general and local effects, and locus-specific effects	Assess the importance of local and general effects using appropriate regressions (Section "Variation in HFC among Loci and the Dichotomy of Local versus General Effects"). Estimate the inbreeding load and its actual impact in the population as in Appendix S1. Do not assume that because only a subset of loci show significant HFCs, they are biologically more relevant than the others. Remember that differences in significance are expected simply as a result of random sampling. If applicable, test whether two groups of loci have a differential effect (e.g., microsatellite vs. allozyme markers, microsatellites in gene-rich and gene-poor regions. . .)—as above, but allowing one regression coefficient for each group, rather than one for each locus.
5. Interpretation of statistical significance	Interpret test statistics cautiously! Significant/nonsignificant estimates do not necessarily imply that a relationship is strong/absent—it only reflects whether a signal happens to cross an arbitrary threshold. Focus on estimates of the actual strength of the signal (r^2, regression slopes, g_2 and their associated standard errors) rather than only on their statistical significance. Do not equate HFC to a measure of selection on identified loci—instead, try to understand what kind of population context (bottlenecks, immigration, nonrandom mating) can enhance or limit HFC in your particular case.

populations (i.e., marine organisms, many invertebrate taxa). Thus, although the construction of robust pedigrees should always be encouraged (Pemberton 2004, 2008), this does not preclude HFC analysis when no other methods are available, or when pedigrees are suspected to poorly reflect true relatedness (Ruiz-Lopez et al. 2009). Importantly, HFC analyses are logistically less demanding than pedigrees—very large populations of small-sized animals can easily be sampled and only one generation is required. Provided there is variance in inbreeding in the population, a few markers ($\leq$10, e.g., Bierne et al. 2000b; Lesbarrères et al. 2007; Townsend et al. 2009) may suffice to reveal inbreeding depression, although large numbers of individuals are needed. Moreover, heterozygosity-based estimates of inbreeding do not require a priori assumptions on relatedness between population members. Indeed, one of the strengths of the HFC approach has been to reveal inbreeding in apparently large, open, populations where its presence had not been anticipated, as in marine bivalves (Singh and Zouros 1978; Koehn and Gaffney 1984; David et al. 1997). Although studies of wild vertebrate populations have been marked for some time by a schism between "pedigree" studies on one side and "molecular" approaches on the other, a recent move toward estimating both pedigree and molecular metrics is welcome and timely (Hedrick et al. 2001; Slate et al. 2004; Overall et al. 2005; Bensch et al. 2006; Jensen et al. 2007; Alho et al.

2009; Ruiz-Lopez et al. 2009, but see Väli et al. 2008). These approaches can be viewed as complementary: pedigrees and HFCs are, respectively, the direct and indirect ways to study inbreeding in populations.

HFC- and pedigree-based inferences have a common limitation: they reveal variation in inbreeding among individuals within the population, but not basic inbreeding shared by all members of the population. However, it is often this "ambient inbreeding" that poses a serious threat to the persistence of populations as it reflects fixed deleterious mutations (i.e., fixation load, as opposed to segregational load). The fixation load can only be assessed by performing artificial crosses among populations (thus measuring heterosis) or correlations between genetic diversity and average population fitness (preferably measured in a "common garden"). The latter generalize the HFC approach to the interpopulation level.

Concluding Remarks

Individual heterozygosity and the inbreeding coefficient are poorly correlated in large panmictic populations at equilibrium. This does not imply that HFCs are not caused by inbreeding depression and that one should invoke alternative explanations such as the presence of major fitness genes in the chromosomal vicinity of marker loci. Particular population structures or histories enhance HFCs by creating correlations in heterozygosity among loci, which may be rather weak (as HFC itself), but extend throughout the genome. The surprise of detecting HFCs, or conversely the disappointment of their absence, expressed in many papers, suggest that natural populations often depart from our initial perception of where HFCs should and should not be expected. It also stems from an insufficient consideration of statistical power issues: the effect size of HFC is usually so low that statistical significance may be a matter of sample size and chance rather than of biology. Before elaborating complicated scenarios to explain variation in HFC, one should remind that significant does not mean strong, and nonsignificant does not mean absent.

HFCs do carry information on inbreeding sensu lato, created by some form of inner structure, or heterogeneity, within populations (e.g., consanguineous vs. nonconsanguineous matings, migrants vs. residents). This in itself warrants credit, as HFCs have revealed the existence of an inner structure and associated fitness costs in populations that were previously thought to be homogeneous, and where there was no other available method to infer this structure. HFC mainly arises when inbreeding created by within-population structure is substantial, and explains a large proportion of the variance in fitness within the population. This property is of prime importance in a conservation genetics context, as HFC can serve as a "warning signal" of genetic erosion in captive or wild populations. However HFC does not yield a detailed understanding of the genetic architecture of animal and plant populations and their response to evolutionary pressures, for which a combination of controlled crossings, molecular genomic tools, and pedigree reconstruction is a more promising research avenue.

ACKNOWLEDGMENTS

This work was financed by the John Fell Fund and an ESF ConGen Short Visit Grant 2488 grant to MS. NB and PD were funded by the Agence Nationale de la Recherche (Hi-Flo project ANR-08-BLAN-0334-01 and SCOBIM project ANR-06-JCJC-02). We sincerely thank L. Keller, J. Radwan, and B.-E. Saether for a valuable discussion, J. Chapman, D. McKenzie, and three anonymous referees for their helpful comments on the manuscript and, G. Leroy for reanalyzing data from Leroy et al. (2009).

LITERATURE CITED

Alho, J. S., B. G. Lillandt, S. Jaari, and J. Merilä. 2009. Multilocus heterozygosity and inbreeding in the Siberian jay. Conserv. Genet. 10:605–609.

Balloux, F., W. Amos, and T. Coulson. 2004. Does heterozygosity estimate inbreeding in real populations? Mol. Ecol. 13:3021–3031.

Bennett, J. H., and F. E. Binet. 1956. Association between Mendelian factors with mixed selfing and random mating. Heredity 10:51–55.

Bensch, S., H. Andren, B. Hansson, H. C. Pedersen, H. Sand, D. Sejberg, P. Wabakken, M. Akesson, and O. Liberg. 2006. Selection for heterozygosity gives hope to a wild population of inbred wolves. Plos ONE 1:e72.

Bierne, N., S. Launey, Y. Naciri-Graven, and F. Bonhomme. 1998. Early effect of inbreeding as revealed by microsatellite analyses on *Ostrea edulis* larvae. Genetics 148:1893–1906.

Bierne, N., A. Tsitrone, and P. David. 2000a. An inbreeding model of associative overdominance during a population bottleneck. Genetics 155:1981–1990.

Bierne, N., I. Bezuart, V. Vonau, F. Bonhomme, and E. Bedier. 2000b. Microsatellite-associated heterosis in hatchery-propagated stocks of the shrimp *Penaeus stylirostris*. Aquaculture 184:203–219.

Britten, H. B. 1996. Meta-analyses of the association between multilocus heterozygosity and fitness. Evolution 50:2158–2164.

Carr, D. E., and M. R. Dudash. 2003. Recent approaches into the genetic basis of inbreeding depression in plants. Philos. Trans. R. Soc. Lond. B 358:1071–1084.

Chakraborty, R. 1981. The distribution of the number of heterozygous loci in an individual in natural populations. Genetics 98:461–466.

———. 1987. Biochemical heterozygosity and phenotypic variability of polygenic traits. Heredity 59:19–28.

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. 1994. The effect of background selection against deleterious mutations on weakly selected, linked variants. Genet. Res. 63:213–227.

Charlesworth, D. 2006. Balancing selection and its effects on sequences in nearby genome regions. Plos Genet. 2:379–384.

Charlesworth, D., and B. Charlesworth. 1987. Inbreeding depression and its evolutionary consequences. Annu. Rev. Ecol. Syst. 18:237–268.

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. Heredity 10:783–796.

Charlesworth, B., M. T. Morgan, and D. Charlesworth. 1991. Multilocus models of inbreeding depression with synergistic selection and partial self-fertilization. Genet. Res. 57:177–194.

Charpentier, M. J. E., F. Prugnolle, O. Gimenez, and A. Widdig. 2008. Genetic heterozygosity and sociality in a primate species. Behav. Genet. 38:151–158.

Coltman, D. W., and J. Slate. 2003. Microsatellite measures of inbreeding: a meta-analysis. Evolution 57:971–983.

Crow, J. F. 2008. Mid-century controversies in population genetics. Annu. Rev. Genet. 42:1–16.

Da Silva, A., G. Luikart, N. G. Yoccoz, A. Cohas, and D. Allaine. 2006. Genetic diversity-fitness correlation revealed by microsatellite analyses in European alpine marmots (*Marmota marmota*). Conserv. Genet. 7:371–382.

Da Silva, A., J. M. Gaillard, N. G. Yoccoz, A. J. M. Hewison, M. Galan, T. Coulson, D. Allaine, L. Vial, D. Delorme, G. Van Laere, et al. 2009. Heterozygosity-fitness correlations revealed by neutral and candidate gene markers in roe deer from a long-term study. Evolution 63:403–417.

Darwin, C. 1876. The effects of cross and self fertilization in the vegetable kingdom. John Murray, London.

David, P. 1997. Modeling the genetic basis of heterosis: tests of alternative hypotheses. Evolution 51:1049–1057.

———. 1998. Heterozygosity-fitness correlations: new perspectives on old problems. Heredity 80:531–537.

David, P., B. Delay, P. Berthou, and P. Jarne. 1995. Alternative models for allozyme-associated heterosis in the marine bivalve *Spisula ovalis*. Genetics 139:1719–1726.

David, P., B. Delay, and P. Jarne. 1997. Heterozygosity and growth in the marine bivalve *Spisula ovalis*: testing alternative hypotheses. Genet. Res. 70:215–223.

David, P., B. Pujol, F. Viard, V. Castella, and J. Goudet. 2007. Reliable selfing rate estimates from imperfect population genetic data. Mol. Ecol. 16:2474–2487.

Eyre-Walker, A. 2006. The genomic rate of adaptive evolution. Trends Ecol. Evol. 21:569–575.

Fu, Y. B., and K. Ritland. 1994. Evidence for the partial dominance of viability genes contributing to inbreeding depression in *Mimulus guttatus*. Genetics 136:323–331.

Gowen, J. W. 1952. Heterosis. Iowa State College Press, Ames.

Haag, C. R., and D. Ebert. 2007. Genotypic selection in *Daphnia* populations consisting of inbred sibships. J. Evol. Biol. 20:881–891.

Hansson, B., and L. Westerberg. 2002. On the correlation between heterozygosity and fitness in natural populations. Mol. Ecol. 11:2467–2474.

———. 2008. Heterozygosity-fitness correlations within inbreeding classes: local or genome-wide effects? Conserv. Genet. 9:73–83.

Hansson, B., S. Bensch, D. Hasselquist, and M. Akesson. 2001. Microsatellite diversity predicts recruitment of sibling great reed warblers. Proc. Soc. B 268:1287–1291.

Hansson, B., H. Westerdahl, D. Hasselquist, M. Akesson, and S. Bensch. 2004. Does linkage disequilibrium generate heterozygosity-fitness correlations in great reed warblers? Evolution 58:870–879.

Hartl, D. L., and A. G. Clark. 2007. Principles of population genetics. Sinauer Associates, Sunderland, MA.

Hedrick, P. 2006a. Genetics of populations. Arizona State Univ., Tempe, AZ.

———. 2006b. Genetic polymorphism in heterogeneous environments: the age of genomics. Annu. Rev. Ecol. Evol. Syst. 37:67–93.

Hedrick, P. W., and O. Muona. 1990. Linkage of viability genes to marker loci in selfing organisms. Heredity 64:67–72.

Hedrick, P., R. Fredrickson, and H. Ellegren. 2001. Evaluation of d-square, a microsatellite measure of inbreeding and outbreeding, in wolves with a known pedigree. Evolution 55:1256–1260.

Hill, W. G., and A. Robertson. 1968. Linkage disequilibrium in finite populations. Theor. Appl. Genet. 38:226–231.

Houle, D. 1989. Allozyme-associated heterosis in *Drosophila melanogaster*. Genetics 123:789–801.

———. 1994. Adaptive distance and the genetic basis of heterosis. Evolution 48:1410–1417.

———. 1998. How should we explain variation in the genetic variance of traits? Genetica 102–3:241–253.

Houle, D., B. Morikawa, and M. Lynch. 1996. Comparing mutational variabilities. Genetics 143:1467–1483.

Hubby, J. L., and R. C. Lewontin. 1966. A molecular approach to the study of genic heterozygosity in natural populations. 1. The number of alleles at different loci in *Drosophila pseudoobscura*. Genetics 54:577–594.

Jacquard, A. 1975. Inbreeding—one word, several meanings. Theor. Popul. Biol. 7:338–363.

Jarne, P., and P. J. L. Lagoda. 1996. Microsatellites, from molecules to populations and back. Trends Ecol. Evol. 11:424–429.

Jensen, H., E. M. Bremset, T. H. Ringsby, and B. E. Saether. 2007. Multilocus heterozygosity and inbreeding depression in an insular house sparrow metapopulation. Mol. Ecol. 16:4066–4078.

Keller, L. F., and D. M. Waller. 2002. Inbreeding effects in wild populations. Trends Ecol. Evol. 17:230–241.

Kimura, M. 1983. The neutral theory of molecular evolution. Cambridge Univ. Press, Cambridge.

Koehn, R. K., and P. M. Gaffney. 1984. Genetic heterozygosity and growth rate in *Mytilus edulis*. Marine Biol. 82:1–7.

Kondrashov, A. S. 1995. Contamination of the genome by very slightly deleterious mutations—why have we not died 100 times over. J. Theor. Biol. 175:583–594.

Kreitman, M. 2000. Methods to detect selection in populations with applications to the human. Annu. Rev. Genomics Hum. Genet. 1:539–559.

Launey, S., and D. Hedgecock. 2001. High genetic load in the Pacific oyster *Crassostrea gigas*. Genetics 159:255–265.

Leroy, G., E. Verrier, J. C. Meriaux, and X. Rognon. 2009. Genetic diversity of dog breeds: within-breed diversity comparing genealogical and molecular data. Anim. Genet. 40:323–332.

Lesbarrères, D., D. S. Schmeller, C. R. Primmer, and J. Merilä. 2007. Genetic variability predicts common frog (*Rana temporaria*) size at metamorphosis in the wild. Heredity 99:41–46.

Lewontin, R. C., and J. L. Hubby. 1966. A molecular approach to the study of genic heterozygosity in natural populations. 2. Amount of variation and degree of heterozygosity in natural populations of *Drosophila pseudoobscura*. Genetics 54:595–609.

Lewontin, R. C., and K. Kojima. 1960. The evolutionary dynamics of complex polymorphisms. Evolution 14:458–472.

Li, M.H, and J. Merilä. 2009. Extensive linkage disequilibrium in a wild bird population. Heredity, *In press*.

Li, H. P., and W. Stephan. 2006. Inferring the demographic history and rate of adaptive substitution in *Drosophila*. Plos. Genet. 2:1580–1589.

Lieutenant-Gosselin, M., and L. Bernatchez. 2006. Local heterozygosity-fitness correlations with global positive effects on fitness in threespine stickleback. Evolution 60:1658–1668.

Lynch, M., and B. Walsh. 1998. Genetics and analysis of quantitative traits. Sinauer Associates, Inc., Sunderland, MA.

Malo, A. F., and T. Coulson. 2009. Heterozygosity-fitness correlations and associative overdominance: new detection method and proof of principle in the Iberian wild boar. Mol. Ecol. 18:2741–2742.

Marshall, T. C., D. W. Coltman, J. Pemberton, J. Slate, J. A. Spalton, F. E. Guinness, J. A. Smith, J. G. Pilkington, and T. H. CluttonBrock. 2002.

Estimating the prevalence of inbreeding from incomplete pedigrees. Proc. R. Soc. Lond. B 269:1533–1539.

Maynard Smith, J., and J. Haigh. 1974. Hitch-hiking effect of a favorable gene. Genet. Res. 23:23–35.

Mitton, J. B., and M. C. Grant. 1984. Associations among protein heterozygosity, growth-rate, and developmental homeostasis. Annu. Rev. Ecol. Syst. 15:479–499.

Nielsen, R., C. Bustamante, A. G. Clark, S. Glanowski, T. B. Sackton, M. J. Hubisz, A. Fledel-Alon, D. M. Tanenbaum, D. Civello, T. J. White, et al. 2005. A scan for positively selected genes in the genomes of humans and chimpanzees. PLoS Biol. 3:976–985.

Ohta, T. 1971. Associative overdominance caused by linked detrimental mutations. Genet. Res. 18:277–286.

———. 1973. Effect of linkage on behavior of mutant genes in finite populations. Theor. Popul. Biol. 4:145–162.

———. 1992. The nearly neutral theory of molecular evolution. Annu. Rev. Ecol. Syst. 23:263–286.

Ohta, T., and C. C. Cockerham. 1974. Detrimental genes with partial selfing and effects on a neutral locus. Genet. Res. 23:191–200.

Ohta, T., and M. Kimura. 1969a. Linkage disequilibrium at steady state determined by random genetic drift and recurrent mutation. Genetics 63:229–238.

———. 1969b. Linkage disequilibrium due to random genetic drift. Genet. Res. 13:47–55.

———. 1970. Development of associative overdominance through linkage disequilibrium in finite populations. Genet. Res. 16:165–177.

Olafsdottir, G. A., and T. Kristjansson. 2008. Correlated pedigree and molecular estimates of inbreeding and their ability to detect inbreeding depression in the Icelandic sheepdog, a recently bottlenecked population of domestic dogs. Conserv. Genet. 9:1639–1641.

Overall, A. D. J., K. A. Byrne, J. G. Pilkington, and J. M. Pemberton. 2005. Heterozygosity, inbreeding and neonatal traits in Soay sheep on St Kilda. Mol. Ecol. 14:3383–3393.

Pamilo, P., and S. Palsson. 1998. Associative overdominance, heterozygosity and fitness. Heredity 81:381–389.

Pemberton, J. M. 2004. Measuring inbreeding depression in the wild: the old ways are the best. Trends Ecol. Evol. 19:613–615.

———. 2008. Wild pedigrees: the way forward. Proc. R. Soc. Lond B 275:613–621.

Pujol, B., P. David, and D. McKey. 2005. Microevolution in agricultural environments: how a traditional Amerindian farming practice favours heterozygosity in cassava (*Manihot esculenta* Crantz, Euphorbiaceae). Ecol. Lett. 8:138–147.

Remington, D. L., and D. M. O'Malley. 2000. Evaluation of major genetic loci contributing to inbreeding depression for survival and early growth in a selfed family of *Pinus taeda*. Evolution 54:1580–1589.

Rousset, F. 2002. Inbreeding and relatedness coefficients: what do they measure? Heredity 88:371–380.

Ruiz-Lopez, M. J., E. R. S. Roldan, G. Espeso, and M. Gomendio. 2009. Pedigrees and microsatellites among endangered ungulates: what do they tell us? Mol. Ecol. 18:1352–1364.

Schaal, B. A., and D. A. Levin. 1976. The demographic genetics of *Liatris cylindracea* Michx. (Compositae). Am. Nat. 110:191–206.

Schierup, M. H., D. Charlesworth, and X. Vekemans. 2000. The effect of hitch-hiking on genes linked to a balanced polymorphism in a subdivided population. Genet. Res. 76:63–73.

Shull, G. H. 1908. The composition of a field of maize. Proc. Am. Breeders Assoc. 4:296–301.

Singh, S. M., and E. Zouros. 1978. Genetic variation associated with growth rate in the American oyster (*Crassostrea virginica*). Evolution 32:342–353.

Slate, J., and J. M. Pemberton. 2007. Admixture and patterns of linkage disequilibrium in a free-living vertebrate population. J. Evol. Biol. 20:1415–1427.

Slate, J., P. David, K. G. Dodds, B. A. Veenvliet, B. C. Glass, T. E. Broad, and J. C. McEwan. 2004. Understanding the relationship between the inbreeding coefficient and multilocus heterozygosity: theoretical expectations and empirical data. Heredity 93:255–265.

Slatkin, M. 1995. Hitchhiking and associative overdominance at a microsatellite locus. Mol. Biol. Evol. 12:473–480.

Smouse, P. E. 1986. The fitness consequences of multiple locus heterozygosity under the multiplicative overdominance and inbreeding depression models. Evolution 40:946–957.

Sved, J. A. 1968. Stability of linked systems of loci with a small population size. Genetics 59:543–563.

Szulkin, M., D. Garant, R. H. McCleery, and B. C. Sheldon. 2007. Inbreeding depression along a life-history continuum in the great tit. J. Evol. Biol. 20:1531–1543.

Townsend, A. K., A. B. Clark, K. J. McGowan, E. L. Buckles, A. D. Miller, and I. J. Lovette. 2009. Disease-mediated inbreeding depression in a large, open population of cooperative crows. Proc. R. Soc. Lond. B 276:2057–2064.

Tsitrone, A., F. Rousset, and P. David. 2001. Heterosis, marker mutational processes and population inbreeding history. Genetics 159:1845–1859.

Väli, U., A. Einarsson, L. Waits, and H. Ellegren. 2008. To what extent do microsatellite markers reflect genome-wide genetic diversity in natural populations? Mol. Ecol. 17:3808–3817.

Vilhunen, S., K. Tiira, A. Laurila, and H. Hirvonen. 2008. The bold and the variable: fish with high heterozygosity act recklessly in the vicinity of predators. Ethology 114:7–15.

Vitalis, R., and D. Couvet. 2001. Two-locus identity probabilities and identity disequilibrium in a partially selfing subdivided population. Genet. Res. 77:67–81.

Weir, B. S., and C. C. Cockerham. 1973. Mixed self and random mating at two loci. Genet. Res. 21:247–262.

Weir, B. S., and W. G. Hill. 1980. Effect of mating structure on variation in linkage disequilibrium. Genetics 95:477–488.

Whitlock, M. 1993. Lack of correlation between heterozygosity and fitness in forked fungus beetles. Heredity 70:574–581.

Wright, S. 1922. Coefficients of inbreeding and relationship. Am. Nat. 56:330–338.

Yang, R. C. 2000. Zygotic associations and multilocus statistics in a nonequilibrium diploid population. Genetics 155:1449–1458.

Zouros, E., and D. W. Foltz. 1987. The use of allelic isozyme variation for the study of heterosis. Isozymes Curr. Topics Biol. Med. Res. 13:1–59.

Zouros, E., S. M. Singh, and H. E. Miles. 1980. Growth rate in oysters: an overdominant phenotype and its possible explanations. Evolution 34:856–867.

Associate Editor: M. Rausher

Supporting Information

The following supporting information is available for this article:

Figure S1. Visualization of identity disequilibrium (g_2) reflected in the distribution of multilocus heterozygosity.
Table S1. Pitfalls in analyzing heterozygosity—fitness data.
Appendix S1. How to estimate the impact of inbreeding on fitness using HFC?
Appendix S2. Disentangling local from general effects when heterozygosities across loci differ markedly.

Supporting Information may be found in the online version of this article.

Please note: Wiley-Blackwell is not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.