Heredity (2004) 92, 306–315
© 2004 Nature Publishing Group All rights reserved 0018-067X/04 $25.00
www.nature.com/hdy

# Neutral locus heterozygosity, inbreeding, and survival in Darwin's ground finches (*Geospiza fortis* and *G. scandens*)

JA Markert[1,2], PR Grant[1], BR Grant[1], LF Keller[3], JL Coombs[4] and K Petren[2]
[1]*Department Ecology and Evolutionary Biology, Princeton University, Princeton, NJ 08544-1003, USA;* [2]*Department Biological Sciences, University of Cincinnati, Cincinnati, OH 45221-0006, USA;* [3]*Zoologisches Museum, Universität Zürich, Winterthurer Str. 190, CH-8057 Zürich, Switzerland;* [4]*Department Neurobiology, Physiology and Behavior, UC Davis, Davis, CA 95616, USA*

Comprehensive long-term studies of isolated populations provide valuable comparative data that may be used to evaluate different methods for quantifying the relationship between genetic diversity and fitness. Here, we report on data collected from large and well-characterized cohorts of the two numerically dominant species of Darwin's finches on Isla Daphne Major, Galápagos, Ecuador – *Geospiza fortis* and *G. scandens*. Multilocus microsatellite (SSR) genetic diversity estimates (heterozygosity and $\bar{d}^2$) and pedigree-based estimates of the inbreeding coefficient ($f$) were compared to each other and to two fitness components: lifespan and recruitment. In the larger sample of *G. fortis*, heterozygosity ($H$) was correlated with both fitness components, but no relationship was detected in the smaller sample of *G. scandens*. Analyses of the inbreeding coefficient detected highly significant relationships between $f$ and recruitment, but no relationship between $f$ and overall lifespan. The $\bar{d}^2$ statistic showed no relationship to either fitness component. When the two SSR-based estimators were compared to $f$, $\bar{d}^2$ was correlated with $f$ in *G. fortis* in the predicted direction, while in *G. scandens* the relationship was positive. Multilocus heterozygosity was correlated with $f$ in *G. fortis* but not in the *G. scandens* sample. A pedigree simulation demonstrated that the variation in true autozygosity can be large among individuals with the same level of inbreeding. This observation may supplement the interpretation of patterns relevant to the local (locus-specific) and general (genome-wide) effects hypotheses, which have been proposed to explain the mechanism responsible for associations between genetic diversity and fitness.
*Heredity* (2004) **92,** 306–315, advance online publication, 21 January 2004; doi:10.1038/sj.hdy.6800409

**Keywords:** *Geospiza*; inbreeding; heterozygosity; heterosis; Darwin's finches; general effects

## Introduction

The loss of genetic diversity and inbreeding depression have been identified as the major potential intrinsic threats to the health and survival of small populations (Frankham, 1998; Vucetich and Waite, 1999). Small and isolated populations are vulnerable to fitness decreases resulting from a loss of genetic diversity, whether this loss results from increased mating among relatives or, on a longer time scale, through genetic drift and fixation of deleterious genotypes (Hedrick and Kalinowski, 2000; Hedrick, 2001). Reduced allelic diversity in a population may be harmful in that it leads to a decrease in individual heterozygosity and reduced reproduction or survival. Two genetic mechanisms can account for the short-term fitness consequences of decreased heterozygosity of populations (and constituent individuals); (1) heterosis, where heterozygous genotypes are superior to any of the homozygous genotypes, and (2) partial dominance, where decreased allelic diversity leads to the expression of recessive or partially recessive deleterious alleles (Houle, 1994). A longer-term consequence is that the loss of alleles from a population reduces a population's capacity to adapt to changing environmental conditions (Lande, 1995; Frankham *et al*, 1999).

The relative levels of genetic diversity within populations and within individuals are of interest, but the most practical and effective method of quantifying such variation is not obvious. Nor is it clear as to which index of genetic diversity is the best predictor of fitness-related traits. When heterozygosity influences fitness, the important quantity to determine is autozygosity, or the total proportion of an individual's alleles that are identical by descent. The inbreeding coefficient ($f$) is one way to estimate this quantity, and it may be determined by analyzing pedigrees (see review by Keller and Waller, 2002). Alternatively, autozygosity may be estimated through the use of biochemical or molecular genetic markers. When allele sizes at microsatellite or simple sequence repetitive (SSR) markers are used, additional statistics that integrate information on the putatively predominant stepwise mutation process may also reflect relatedness (Coulson *et al*, 1998). If they perform well, molecular alternatives to $f$ are desirable because measuring the extent of inbreeding through

*Correspondence: J Markert, Department Biological Sciences, University of Cincinnati, Cincinnati, OH 45221-0006, USA.*
*E-mail: Jeffrey.Markert@UC.edu*
Received 29 July 2003; accepted 10 October 2003

pedigrees in natural or disturbed populations is often difficult because it requires repeated sampling and observation over several generations.

Despite these difficulties, several studies have used the inbreeding coefficient to document inbreeding depression in populations that are small, isolated, and well enough studied for extensive pedigrees to be ascertained. In a recent review, Keller and Waller (2002) identified 18 animal species where pedigrees were used to study inbreeding in natural populations. Not surprisingly, 12 of the species were birds and six were large, long-lived mammals. In these cases, individuals could be tagged or otherwise recognized. Where such intensive and long-term sampling is not possible, snapshot methods using molecular or biochemical markers may be the only tools available to estimate the extent of inbreeding.

Correlations of heterozygosity and fitness have been examined in a variety of organisms using both allozyme and DNA markers. Study organisms include marine bivalves, trees, crustaceans, amphibians, birds, and mammals (Zouros and Foltz, 1987; Britten, 1996; Bierne *et al*, 1998; Westemeier *et al*, 1998; Bierne *et al*, 2000; Launey and Hedgecock, 2001; Myrand *et al*, 2002). It is usually assumed that individuals (or populations) with reduced heterozygosity are more inbred than related populations with greater heterozygosity. Reduced individual or population heterozygosity at a set of marker loci is often, but not always (Booth *et al*, 1990; Savolainen and Hedrick, 1995; Rowe and Beebee, 2001) associated with a decrease in some fitness-related trait. In addition, there are examples of genetically depauperate populations being 'rescued' by the introduction of new alleles (Westemeier *et al*, 1998; Grant *et al*, 2001; Vila *et al*, 2002).

Heterozygosity–fitness correlations (HFC's) are often very weak, and their biological relevance has been questioned. David (1998), for example, has asked whether the published studies that demonstrate small positive HFCs might simply reflect a publication bias in favor of positive results (for a discussion, see Coltman and Slate, 2003). It is difficult to generalize about the prevalence of HFC's from individual studies because the strength of HFC's and inbreeding depression may vary with environmental conditions (David and Jarne, 1997; Keller *et al*, 2002; Knaepkens *et al*, 2002; Myrand *et al*, 2002).

Since heterozygosity can sometimes predict fitness even among siblings or other sets of identically inbred individuals, it has been suggested that linkage between marker and coding loci is responsible for HFC's, a pattern termed 'local effects' by David *et al* (1995). For example, in great reed warblers, more heterozygous siblings are more likely to become recruits (Hansson *et al*, 2001), and broods of the oyster *Ostrea edulis* become increasingly heterozygous over time (Bierne *et al*, 1998). Recently, Hansson and Westerberg (2002) have noted that data sets that include information on $f$, $H$, and fitness are especially helpful in differentiating between the local effects hypothesis and the alternative 'general effects' hypothesis, which suggests that HFC's are due to genome-wide heterozygosity, and that both $H$ and $f$ act as different but potentially complementary indices of autozygosity (David, 1998).

In a multi year study on Daphne Major, Keller *et al* (2002) have used pedigree data to show that inbreeding depression is present in both *G. scandens* and *G. fortis*. The severity of the effect is related to annual rainfall, which in turn influences food availability. In this study, we focus on a single generation of the birds studied by Keller *et al* (2002), which allows us to deepen our understanding of the relationship between pedigree estimates of inbreeding and heterozygosity by directly comparing the two estimators, and in turn determining how well each of these estimators predicts fitness. We examine the relationships between genetic diversity and fitness by examining cohorts of birds hatched in 1991 in two species of Darwin's finches, the medium ground finch *Geospiza fortis* and the cactus finch *G. scandens*. We compare microsatellite heterozygosity, the stepwise mutation model-based $\bar{d}^2$ statistic, and a pedigree-derived estimate of the inbreeding coefficient taken from Keller *et al* (2002) as predictors of lifespan and of recruitment into the next generation. These two cohorts were selected because they represent the first year in which DNA samples are available for the majority of individuals hatched in a long-term study on Isla Daphne Major (Grant and Grant, 2002). Following a single cohort has the advantage of comparing sets of individuals who have shared a common environment since birth, thereby eliminating differential effects of environmental variation during their lifetimes. Using this unique data set, we are able to explore the properties of these estimators in a natural population.

## Methods

### Field methods

Isla Daphne Major is a small (0. 34 km$^2$) island about 7.5 km from the nearest larger island of Santa Cruz. Almost every bird hatched in 1991 was banded either as a nestling, or as an adult (Grant and Grant, 2002 and references therein). Blood samples were obtained from the two numerically dominant species, *G. fortis* and *G. scandens* via brachial vein puncture usually from 8-day-old nestlings, treated with EDTA and dried on filter paper (Petren *et al*, 1999b). Hatching dates, social parents, and sibships were recorded during the field season. *G. fortis* hatched between mid-March and mid-June 1991 and *G. scandens* hatched between mid-January and early June. Lifespan and recruitment were chosen as fitness correlates because they could be unambiguously determined in the field, and each has a direct relationship to fitness (Grant and Grant, 2000). Lifespan was determined during annual surveys of marked birds each spring. For the purposes of this study, recruitment was determined by whether or not an individual was observed attempting to breed during its lifetime.

### Laboratory

Microsatellite genotypes at 13 loci were obtained for 470 of the 782 *G. fortis* and 110 of the 138 *G. scandens* that hatched during 1991. Fragment sizes were obtained using previously reported PCR methods for the following presumably unlinked loci that were developed from a *G. fortis* library (Petren, 1998): Gf-1, Gf-3, Gf-4, Gf-5, Gf-7, Gf-8, Gf-9, Gf11, Gf-12, Gf-13, Gf14, Gf-15, and Gf-16. Genotypes were determined using the methods described in Petren (1998). Fragment sizes were determined on an ABI 377 DNA sequencer, an ABI 3100, or by autoradiography.

### Data analysis

Standard indices of genetic diversity were calculated from microsatellite genotypes for each species and individual, including unbiased heterozygosity ($H$) (Nei, 1978), the effective number of alleles ($n_e$) (Hartl and Clark, 1989), and a count of the number of alleles present. A stepwise mutation model-based estimator of genetic diversity, $\bar{d}^2$ (Coulson *et al*, 1998), was also calculated and is included here for comparative purposes. For many of the individuals in this current study, pedigree-based inbreeding coefficients were taken from Keller *et al* (2002). To reduce the error that could be introduced by using $f$ values from birds with incompletely defined pedigrees, we used a restricted data set for making comparisons involving $f$. We included only those individuals with four known grandparents and whose fathers could be confirmed using microsatellite genotypes. In these species, maternity can be reliably inferred; however, 20% of *G. fortis* (Keller *et al*, 2001) and 8% of *G. scandens* (Petren *et al*, 1999a) are not genetically related to their social father. Restricted samples of 211 *G. fortis* and 75 *G. scandens* were available for comparisons between $f$ and either $H$ or $\bar{d}^2$ and fitness-related traits. Correlations between the inbreeding coefficient and either $d^2$ or $H$ were determined using Spearman's rank correlation ($r_s$) because $f$ is heavily biased towards zero in these samples. In contrast, the distributions of $H$ and $\bar{d}^2$, while not normally distributed, were less biased and were analyzed using parametric statistics.

We used bootstrap resampling of individuals from the large *G. fortis* sample to determine the effects of sample size differences on correlations between $H$ and lifespan. In addition, 95% confidence intervals were calculated for many parameters to determine whether the failure to reject the null was likely to be a consequence of inadequate sample sizes (Hoenig and Heisey, 2001). In these analyses, family structure may complicate the interpretation of results because broods are identical in inbreeding and because fitness may be influenced by parental effects. However, we analyze individuals because our goal is to compare fitness, heterozygosity, and autozygosity due to inbreeding and these are individual properties that may vary among identically inbred individuals (Hansson and Westerberg, 2002).

The mean values of $f$, $H$, and $\bar{d}^2$ were compared in recruits and nonrecruits, where nonrecruits were defined as birds who were never observed attempting to breed, and includes birds who died before sexual maturity. The Mann–Whitney $U$ test was used to analyze inbreeding data, while $t$-tests were used to evaluate $H$ and $\bar{d}^2$ that are closer to being normally distributed.

To isolate the effects of heterozygosity from those of inbreeding, we compared $H$ and fitness among individuals with the same $f$ value in cases where more than 10 individuals shared the same inbreeding coefficient.

Finally, to understand the expected level of variation in genetic similarity between an individual's chromosomes for a given level of inbreeding, a series of pedigrees were simulated to generate individuals with theoretical inbreeding coefficients of 0.25, 0.125, and 0.0625, and the number of linkage groups was arbitrarily assumed to be either 15, 40, or 60 in order to determine the effect of linkage on variance in autozygosity. The starting generation consisted of two individuals that possessed four different alleles at each locus, so that the starting heterozygosity was always 100%. In total, 1000 replicates for each of the nine conditions were generated. This permitted the calculation of 95% confidence intervals (the central 95% of synthetic individuals) for the mean decrease in heterozygosity expected among individuals for the three levels of inbreeding under conditions of no linkage.

## Results

### Allelic diversity

In *G. fortis*, 150 unique alleles were observed at the 13 SSR loci surveyed, while in the smaller sample of *G. scandens* 114 alleles were observed, for an average of 11.5 and 8.8 alleles, respectively. The 13 locus average effective number of alleles ($n_e$) – the number of equally frequent alleles that would produce the expected heterozygosity derived from the actual allele frequencies (Hartl and Clark, 1989) – is 4.26 in *G. fortis*, and ranges from a low of 1.79 at locus Gf 4 to a high of 6.29 at locus Gf 11. In *G. scandens*, $n_e$ is 3.88 when averaged over all 13 loci and ranges from a low of 1.88 at locus Gf 4 to a high of 7.42 at locus Gf 12. Values for $\bar{d}^2$ ranged from 9.1 to 488.2 in *G. fortis*, with an average of 202.98 and from 18.2 to 496 in *G. scandens* with an average of 155.53. Observed heterozygosity is 0.65 in *G. fortis* with a SD of 0.14 and is 0.61 with a SD of 0.13 in *G. scandens*.

### Inbreeding

The inbreeding coefficient could be reliably determined for 211 *G. fortis* and 75 *G. scandens*. In *G. fortis*, 33% of these individuals have a pedigree-based inbreeding coefficient greater than 0, while in *G. scandens* 48% of this subset had an $f$ greater than 0. This fraction of detectably inbred birds in both species is higher than that reported in the multigeneration study of Keller *et al* (2002), who observed that 19.8% of *G. fortis* and 16.7% of *G. scandens* were detectably inbred. The mean level of inbreeding observed in this study is $f = 0.010$ in *G. fortis* and $f = 0.042$ in *G. scandens*.

### Inbreeding and genetic diversity

To test the relationship between the inbreeding coefficient and heterozygosity, Spearman's Rho ($r_s$) was calculated. Heterozygosity and $f$ were negatively correlated in *G. fortis* ($r_s = -0.162$, $P = 0.019$) and not correlated in *G. scandens* ($r_s = 0.087$, $P = 0.45$). The mean heterozygosity and the number of individuals in the sample for each inbreeding class are shown in Figure 1. Several inbreeding classes are consistent with the expected $H$ in *G. fortis*, whereas all of the inbreeding classes $> f = 0$ are above expectation in *G. scandens*. Many inbreeding classes contained small numbers of individuals in both species. In *G. scandens*, many individuals are also outside the 95% confidence interval for autozygosity derived from the pedigree simulation.

When the $\bar{d}^2$ statistic was substituted for $H$, an even stronger negative correlation was observed in *G. fortis* ($r_s = -0.224$, $P = 0.0012$). Surprisingly, a significant *positive* correlation was observed between $f$ and $\bar{d}^2$ in *G. scandens* ($r_s = 0.307$, $P = 0.0083$). Thus, both diversity indices are consistent with the basic prediction that genetic diversity will decrease with increased inbreeding in *G. fortis*, but in *G. scandens* the relationship between $f$

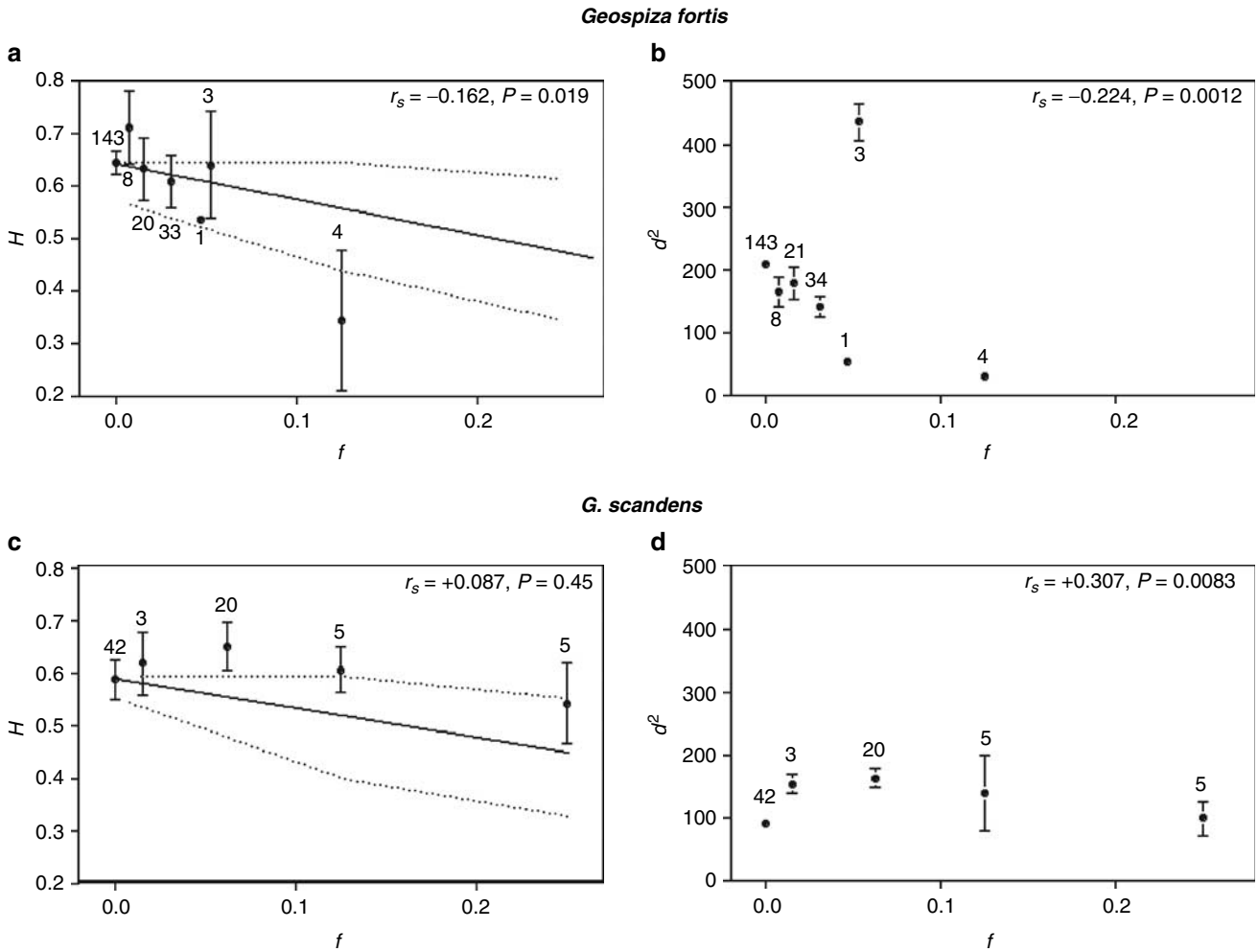


**Figure 1** Mean inbreeding *versus* $H$ or $\bar{d}^2$ in *G. fortis* (**a** and **b**) and *G. scandens* (**c** and **d**). Mean values for a given level of inbreeding are shown (closed circles). The number of individuals observed for each level of inbreeding is shown in the figure, as are the Spearman rank correlations ($r_s$) for each comparison. The black lines in panels a and c are the expected decline in $H$ calculated from the theoretical increase in autozygosity for each inbreeding class. The broken lines on these panels represent the 95% confidence intervals for predicted decline in heterozygosity at 15 unlinked loci based on a simulation of 1000 pedigrees.

and $H$ may be undetectable for smaller sample sizes, and the $\bar{d}^2$ statistic may even be misleading.

## Fitness and genetic diversity

Correlations between either $f$, $H$, or $\bar{d}^2$ and lifespan were calculated to determine how well the three indices of genetic diversity predicted fitness. Lifespan and $H$ were correlated in *G. fortis* ($r = 0.17$, $P = 0.001$), but not in *G. scandens*. Neither $f$ nor $\bar{d}^2$ were correlated with lifespan in either taxon. The results are shown in Table 1, along with 95% confidence intervals for $r$. Recruits had higher average $H$ and $\bar{d}^2$ values than nonrecruits in both *G. fortis* and *G. scandens*; however, the only significant difference involved $H$ in *G. fortis* ($H = 0.68$ *vs* 0.63, $t_{468} = -4.3$, $P < 0.0001$). The Mann–Whitney $U$ test results for $f$ show that recruits were on average significantly less inbred than nonrecruits in both taxa (Table 2). Tsitrone *et al* (2001) have shown that the predicted relationship between fitness and $\bar{d}^2$ is nonlinear, and our results also suggest a curvilinear relationship.

Average $H$ or average $f$ of survivors and nonsurvivors was plotted for each year to reveal when genetic diversity influences survival most strongly (Figure 2). In the *G. fortis* cohort, the inbreeding coefficient decreased and heterozygosity increased steadily during the first 3 years as more inbred and less heterozygous individuals died more frequently. The differential mortality appears to be most strongly expressed during the first 2 years of life. Survivors (closed squares) are clearly distinguishable from nonsurvivors (open circles) with respect to $H$ during 1991 and 1992 in *G. fortis* ($H = 0.66$ and 0.62, $t_{468} = -3.33$, $P = 0.0009$ for 1 year survivors and nonsurvivors, respectively), and become more similar in subsequent years. Similarly, the mean inbreeding coefficient is higher in individuals who died during 1991, 1992, or 1993, although there is substantial overlap in the distributions. The relationship changes in later years when the number of individuals surviving in the sample becomes smaller. Survivors and nonsurvivors were essentially indistinguishable in *G. scandens*, which is not surprising given the lack of an overall correlation between genetic diversity and either lifespan or recruitment.

Sample size limitations affect our ability to detect temporal patterns of variation in fitness and correlations

between genetic diversity indices and fitness, especially in *G. scandens*. To investigate sample size effects on correlations, we drew at random subsamples of 100 individuals 100 times from the larger *G. fortis* data set. In all, 40% of these subsamples showed a significant correlation between lifespan and *H*, suggesting there is roughly a 60% chance of a type II error in *G. scandens* if they actually have a HFC similar in magnitude to that observed in *G. fortis*. An alternative comparison can be made using the reasoning of Hoenig and Heisey (2001). They suggest that when 95% confidence intervals are clustered tightly around the null (in this case $r=0$), then failure to reject the null is not likely due to small sample sizes. The confidence interval of *r* ranged from −0.14 to 0.2 in *G. scandens*, which is a rather broad distribution around the null given the modest size of the effect observed in *G. fortis* ($r=0.16$). This suggests that the null might be rejected if sample size were expanded. A similar pattern was observed for the nonsignificant difference between *H* in recruits and nonrecruits in *G. scandens* (Table 2).

A more qualitative evaluation of the *G. fortis* data shows that the small number of individuals in the three most inbred classes have a decrease in lifespan, and the most heterozygous individuals live longer than the least heterozygous individuals (Figure 3a**)**. In contrast, the fitness effects of inbreeding are less evident in *G. scandens*, although the most inbred birds survive less well (Figure 3d**)**. Similarly, no clear pattern or hint of a trend emerges from an examination of the relative lifespans of the different *H* classes in *G. scandens* (Figure 3c**)**.

The local effects hypothesis predicts that HFC's should be detectable among individuals with the same level of inbreeding (David, 1998). In *G. fortis*, three levels of inbreeding were represented by more than 10 individuals (an arbitrarily chosen minimum); $f=0$, 0.031 and 0.156. In *G. scandens*, only $f=0$ and 0.0625 were represented by more than 10 individuals. No correlations between *H* and lifespan were detected among the individuals within these inbreeding classes. A difference in mean *H* was detected between recruits and nonrecruits in *G. fortis* with an inbreeding coefficient of 0.031 where *H* in recruits was 0.69 compared to 0.56 in nonrecruits ($N=33$, $t_{31}=-2.489$, $P=0.018$), although this difference becomes marginally insignificant after correcting for three comparisons ($P=0.056$; Table 3**)**. However, all of the relationships are in the direction predicted by the local effects hypothesis, and the 95% confidence intervals suggest a lack of statistical power due to small sample sizes.

Variation in the true level of autozygosity predicted by *f* is a potentially important source of noise in comparisons of inbreeding and fitness. Pedigree simulations revealed that mean values of autozygosity for groups of individuals quickly converged on that predicted by *f*; however, the variance in the expected decline in *H* among individuals is high for all levels of inbreeding and for each number of simulated linkage groups. For example, the 95% confidence interval for heterozygosity decline in individual offspring of full-sib matings ($f=0.25$) ranges from 6.7 to 47% when heterozygosity is sampled at 15 unlinked loci. With 60 loci this range is smaller, but still large (from 15 to 37%) (Table 4), suggesting that there is a substantial amount of autozygosity variance that is not accounted for by *f*. This variance may affect the ability to detect fitness associations when modest numbers of individuals are sampled or when the amount of linkage disequilibrium in a population is high as linkage would be expected to further increase the variance in autozygosity (Franklin, 1977). To show the power to detect correlations between *H* and *f*, we plotted confidence intervals for 15 loci (Figure 1).

**Table 1** Correlation coefficients between *f*, *H*, or $\bar{d}^2$ and lifespan in *G. fortis* and *G. scandens*

| | *G. fortis* | *G. scandens* |
|---|---|---|
| *f versus* lifespan – $r_s$ (*P*) | 0.004 (0.36) | 0.006 (0.9613) |
| *H versus* lifespan: *r* | 0.16 (<0.001) | 0.03 (0.8051) |
| (*P*)[95% ci] | [0.1,0.24] | [−0.14,0.2] |
| $\bar{d}^2$ *versus* lifespan: *r* | 0.02 (0.6388) | 0.07 (0.4462) |
| (*P*)[95% ci] | [−0.1,0.1] | [−0.1,0.26] |

*P*-values are shown in parentheses, and 95% CIs for *r* are shown in square brackets. Spearman's rank correlations were calculated for comparisons involving *f*.

## Discussion

Comprehensive long-term studies of isolated populations provide valuable comparative data that may be used to evaluate different approaches to detect HFCs. Using the current Darwin's finch data set, we addressed two distinct kinds of questions; (1) those involving the relationship between the inbreeding coefficient (*f*) and indices of genetic diversity derived from SSR loci; and (2)

**Table 2** Comparisions of genetic diversity between recruits and nonrecruits

| | *G. fortis* | | | *G. scandens* | | |
|---|---|---|---|---|---|---|
| | f | H | $\bar{d}^2$ | f | H | $\bar{d}^2$ |
| Recruits | 0.007 | 0.68 | 207.6 | 0.028 | 0.6 | 149.8 |
| Nonrecruits | 0.011 | 0.63 | 200.6 | 0.045 | 0.62 | 128.5 |
| Difference (95% CI) | −0.004 | 0.056(0.032,0.081) | 7(−29.9,38.4) | −0.017 | 0.087(−0.021,0.196) | 21.3(−64.3,21.7) |
| Rank recruits, nonrecurits | 120.7,93.7 | — | — | 66.6,33.1 | — | — |
| *t* | — | −4.3 | −0.639 | — | −0.036 | −0.98 |
| *P*-value | 0.012 | <0.0001 | 0.5232 | <0.0001 | 0.96 | 0.329 |

Mann–Whitney *U* tests were used to compare the ranks of recruits and nonrecruits with respect to *f*. Heterozygosity and $\bar{d}^2$ differences were evaluated using *t*-tests. Ranks are shown for recruits and nonrecruits for comparisons involving *f*. The 95% confidence intervals are shown for the difference in the mean for comparisons using parametric statistics.

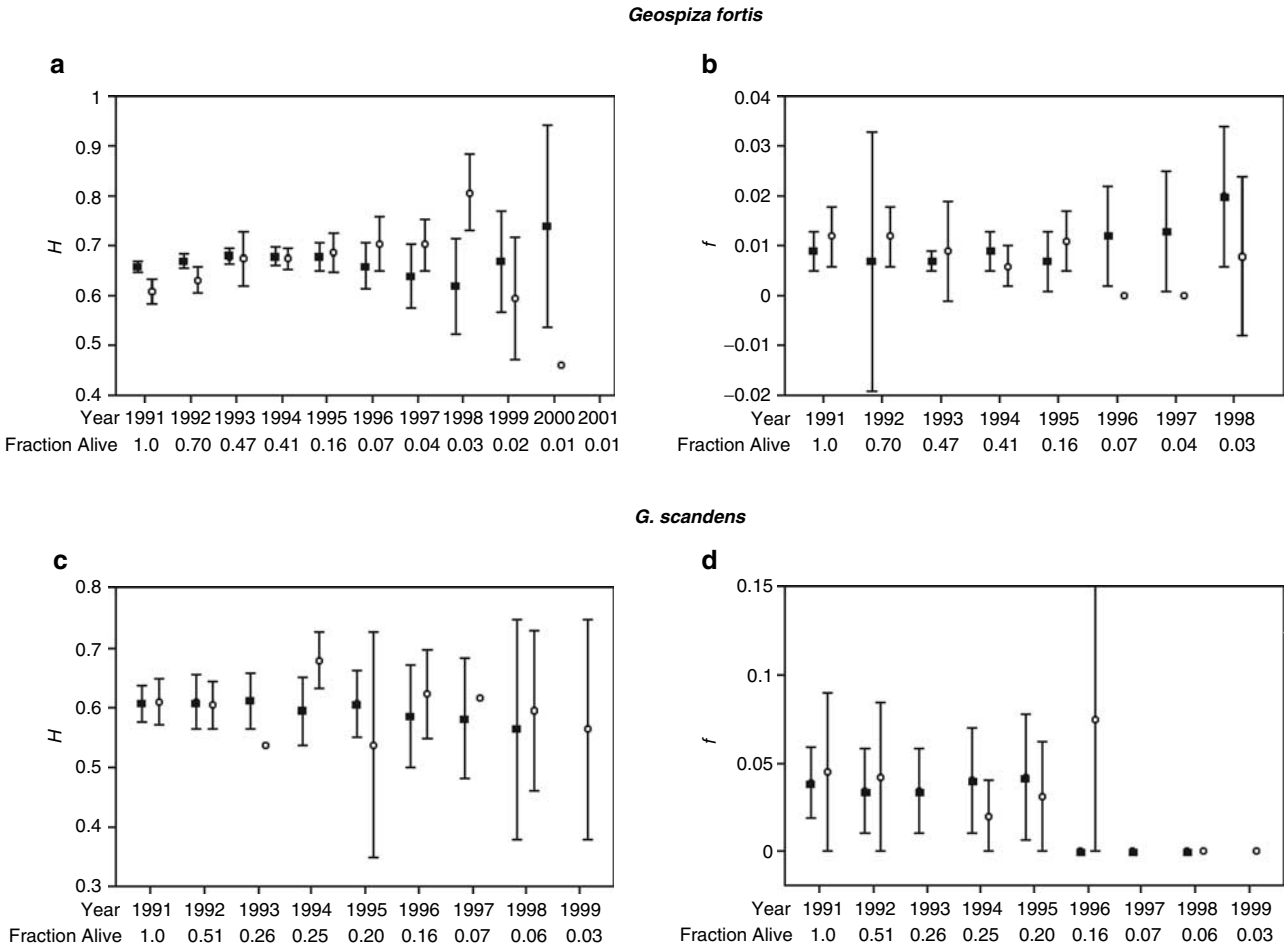


**Figure 2** A year by year comparison of mean heterozygosity ($\pm 2$ SE) for individuals who die during a year (open circles) or who survive until at least the next year (closed squares). *G. fortis* comparisons are shown in panels **a** and **b**, and *G. scandens* comparisons are shown in panels **c** and **d**. The fraction of the total sample living at the beginning of an interval is shown below each graph. Starting sample sizes in *G. fortis* are 470 individuals for *H* and 211 individuals for *f*. In *G. scandens*, starting sample sizes are 110 for *H* and 75 for *f*. Confidence intervals of 0 reflect identical *f* or *H* values within a subgroup.

those concerning the relative abilities of *f* and the SSR-derived estimators of autozygosity to predict fitness declines resulting from inbreeding. These results underscore the fact that both molecular estimates of autozygosity ($H$ & $\bar{d}^2$) and the inbreeding coefficient calculated from pedigrees ($f$) are imperfect estimators of expected individual autozygosity, a quantity that represents the true proportion of loci which have the potential to show heterotic effects. In summary, we find a correlation between *f* and both *H* and $\bar{d}^2$ in the larger sample of *G. fortis*, but heterozygosity is not correlated with inbreeding in the smaller *G. scandens* sample. The correlation of the $\bar{d}^2$ statistic with *f* in *G. scandens* was opposite to the predicted direction, suggesting that *H* is a more reliable index of SSR diversity than $\bar{d}^2$, a finding consistent with recent studies of the behavior of this estimator (reviewed by Goudet and Keller, 2002). With respect to fitness, *H* is correlated with lifespan in *G. fortis*, but not in *G. scandens*. The inbreeding coefficient is significantly higher in nonrecruits in both species sampled, whereas *H* is significantly higher in recruits only in the larger sample of *G. fortis*. In *G. fortis*, the most pronounced effect of *H* on lifespan was detected during the first 2 years life. A closer examination of these patterns can illuminate the behavior of these estimators in natural populations.

These results may be compared with those of Keller *et al* (2002). In contrast to the results presented here, Keller *et al* (2002) found the strongest evidence of inbreeding depression in *G. scandens*. One explanation for this difference is that the expression of inbreeding depression varies from season to season in response to rainfall (Keller *et al*, 2002). The 1991 cohorts studied here hatched during the first of 3 wet years (Grant *et al*, 2000), which would likely have produced a different selective environment as the majority of individuals surveyed in Keller's *et al* (2002) multiyear study experienced much drier conditions.

## Predicted heterozygosity and inbreeding

In *G. fortis*, the mean levels of observed heterozygosity at the 13 SSR loci are close to the values expected for a given decline in autozygosity for most levels of inbreeding, although the variation around this mean is large. In contrast, the observed heterozygosities in the smaller *G. scandens* sample are not close to those predicted. One potential source of this variation involves the fact that

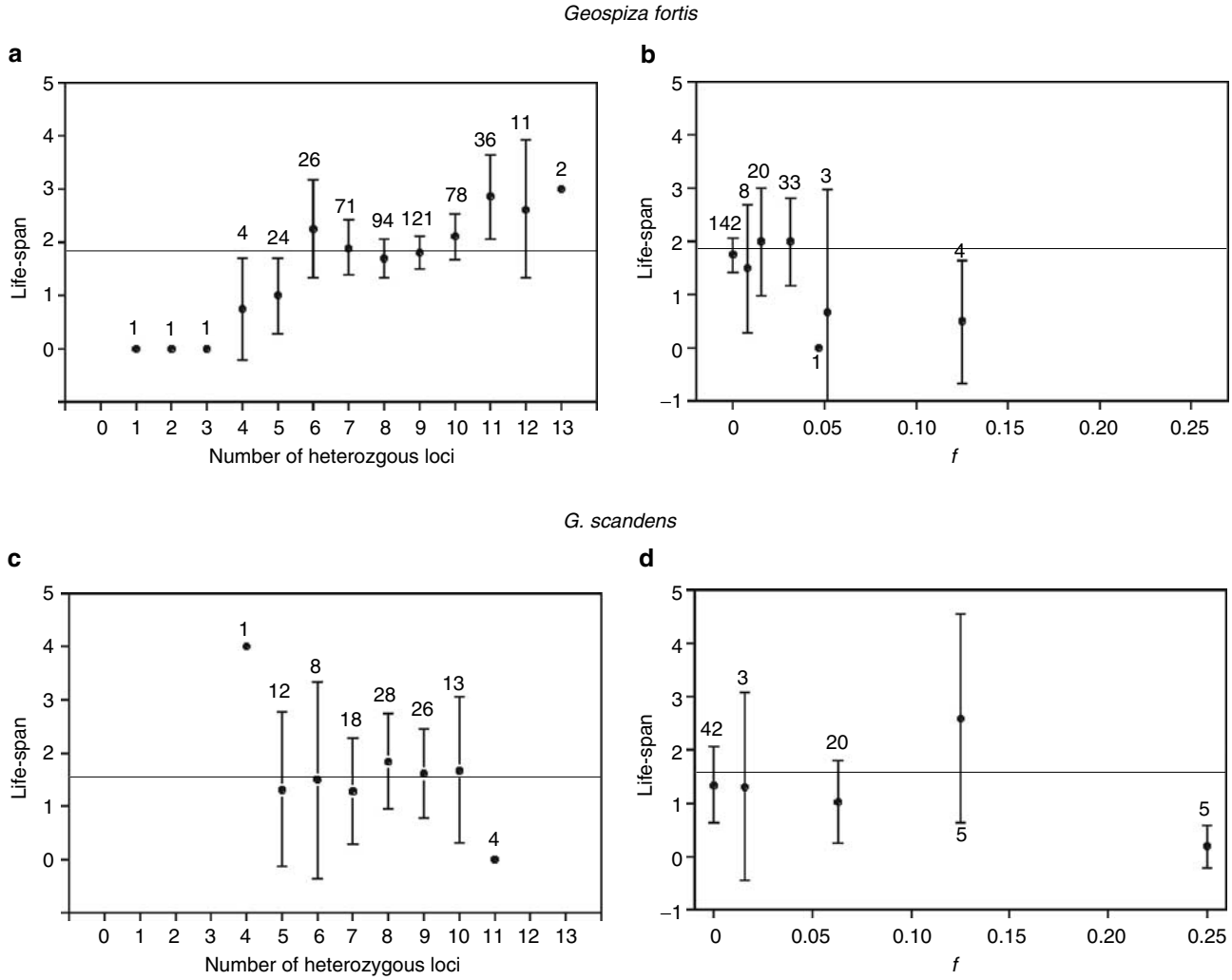


**Figure 3** Indices of autozygosity as predictors of fitness showing 95% confidence intervals for lifespan for each observed level of either *H* or *f*. The number of individuals observed in each class is shown on the graph. The light gray line indicates mean lifespan for the sample. *G. fortis* results are shown in panels **a** and **b**, *G. scandens* results are shown in panels **c** and **d**. Confidence intervals of 0 reflect identical *f* or *H* values within a subgroup.

pedigree-based estimates of *f* are calculated relative to a starting population whose members must be assumed to be unrelated. When this occurs, variation in autozygosity due to undetected inbreeding differences in this reference population leads to heterogeneity in the true level of autozygosity among individuals who share an inbreeding coefficient, although the influence of this variation is expected to decrease with pedigree depth (Keller and Waller, 2002). A second possibility is that increased mortality of homozygous *G. scandens* prior to blood sampling at day 8 could cause a similar pattern.

A third source of variation involves the sampling of chromosomal regions during meiosis. If linkage groups are large, then variance in autozygosity for a given level of inbreeding is expected to be large (Franklin, 1977). If so, then variance in the true level of autozygosity may account for a large part of the discrepancy between observed and expected heterozygosity in *G. scandens*, as much of the observed individual heterozygosity values fall within the 95% confidence intervals for autozygosity. This variation in autozygosity is an important influence on the number of heterozygous coding loci that are visible to selection. Variation in actual identity by descent among loci is expected to be high with a sample of 13 unlinked marker loci, as in this study. The results from the simulation exercise allow us to conclude that this kind of variation leads to considerable heterogeneity in autozygosity within an inbreeding class. The effects of this heterogeneity on the relationships between *f* and *H* will be most pronounced when either sample size or the true number of linkage groups is modest; however, the variance is considerable even with 60 unlinked loci. Of course the strength of the correlation between *H* and fitness is also prone to other sources of error such as environmental variation and parental effects, so it is not surprising that HFC's are typically weak (David, 1998).

## Genetic diversity and fitness

From these data, it is unclear whether the inbreeding coefficient or heterozygosity is a better predictor of fitness. Ideally, multiple regression would be used to directly compare the relative abilities of SSR heterozygosity and the inbreeding coefficient to predict fitness.

**Table 3** A comparison of fitness with heterozygosity within levels of inbreeding

| *Level of inbreeding* | *0* | *0.0156* | *0.031* | *0.0625* |
|---|---|---|---|---|
| Number of Obs. | 142 | 20 | 33 | |
| *r*: *H versus* lifespan (95% CI) | 0.15 (0,0.3) | 0.33 (−0.14,0.66) | 0.16 (−0.17,0.48) | |
| *F* ratio | 3.13 | 2.25 | 0.904 | |
| *P* | 0.07 | 0.15 | 0.35 | |
| Mean *H:* recruits | 0.67 (0.63,0.71) | 0.68 (0.57,0.78) | 0.69 (0.61,0.78) | |
| Mean *H:* nonrecruits | 0.64 (0.61,0.66) | 0.59 (0.52,0.67) | 0.56 (0.51,0.62) | |
| Difference: recruits/nonrecruits (95% CI) | 0.056 (0.032,0.081) | 0.084 (−0.047,0.21) | 0.13 (0.024,0.23) | |
| $t_{n-2}$ | −1.518 | −1.35 | −2.489 | |
| *P* | 0.13 | 0.19 | 0.018 | |
| Number of Obs. | 45 | | | 21 |
| *r*: *H versus* lifespan | 0.18 (−0.14,0.46) | | | 0.21 (−0.26,0.59) |
| *F* ratio | 1.37 | | | 0.6 |
| *P* | 0.25 | | | 0.381 |
| Mean *H:* recruits | 0.61 (0.51,0.70) | | | 0.69 |
| Mean *H:* nonrecruits | 0.59 (0.54,0.63) | | | 0.65 (0.59,0.70) |
| Difference: recruits/nonrecruits (95% CI) | 0.018 (−0.021,0.196) | | | — |
| $t_{n-2}$ | −0.345 | | | — |
| *P* | 0.72 | | | — |

The 95% confidence intervals are shown for the correlation between *H* and lifespan and for the difference between mean *H* in recruits *versus* nonrecruits. Only inbreeding classes where >10 individuals were observed were analyzed. In *G. scandens,* for $f=0.0625$, a *P*-value was not calculated for recruitment as only a single individual was recruited.

**Table 4** Pedigree resampling results showing 95% confidence intervals for the expected decline in heterozygosity per generation for three levels of inbreeding

| | *Number of linkage groups sampled* | | |
|---|---|---|---|
| *f* | *15* | *40* | *60* |
| 0.0625 | 0.00, 0.20 | 0.00, 0.15 | 0.02, 0.13 |
| 0.125 | 0.00, 0.33 | 0.05, 0.23 | 0.05, 0.21 |
| 0.25 | 0.07, 0.45 | 0.13, 0.38 | 0.15, 0.37 |

In total, 1000 simulated pedigrees were constructed at each of the three levels of inbreeding using either 15, 40, or 60 loci. No alleles were shared among members of the initial generations. The confidence intervals shown represent the central 95% of results from each simulation run.

This is not practical because (as in most studies) the inbreeding coefficient is biased with a median value of $f=0$ and few highly inbred individuals, whereas *H* is closer to being normally distributed. Our analyses show that, in general, more autozygous birds have shorter lives than their outbred contemporaries. In *G. fortis*, this pattern appears to be driven by differential mortality early in life (Figure 2). The mean heterozygosity of *G. fortis* that survive their first year is 4% higher than for birds that do not. Those that survive also have a lower inbreeding coefficient than those that do not. The most inbred *G. scandens* have a reduced life expectancy (Figure 3), although there is no difference in *H* between those who die before their first year and those who live for a year or more. In the case of *G. fortis*, it appears that there is a threshold below which inbreeding has little effect under the benign environmental conditions prevailing early in the life of this particular cohort. Similarly, *G. scandens* with $f=0.0625$ and 0.25 survive less well than individuals with lower levels of inbreeding. An exception to this pattern is the small number of individuals with $f=0.125$, who have a higher survival than average. Three individuals survived less than a year, and the other two (who were siblings) lived 5 years and were recruited into the breeding population. Therefore, only two individuals are having a large influence on the lifespan of this inbreeding class, contradicting the more general pattern and underscoring the perils of small sample sizes. In *G. fortis,* these inbreeding results are paralleled by those for heterozygosity, where the most heterozygous individuals survive longer than the least heterozygous group. No clear relationship emerged when the most heterozygous and homozygous *G. scandens* were compared.

The absence of a correlation between heterozygosity and fitness in *G. scandens* may be due to the small number of individuals in the sample with very high and very low heterozygosities. Hybridizations between *G. scandens* and *G. fortis* are rare events (Grant and Grant, 1992) that provide additional insights into the relationship between heterozygosity and fitness. Grant *et al* (in press) have explored the fitness consequences of these hybridization events in progeny hatched in 1991 on the same island, and found that hybrid (backcross) progeny have higher fitness than nonhybrid contemporaries matched in time and space to minimize the effects of environmental noise. Since hybrid progeny have a higher expected heterozygosity than nonhybrid progeny, this observation suggests that in both species fitness is enhanced by increased heterozygosity.

The limited number of organisms available for study in small populations may be the greatest challenge to those seeking to study biologically important but quantitatively weak effects. In this study, the samples represented a large portion of the entire cohort with genetic data available for 60% of the *G. fortis* and 80% of the *G. scandens* hatched in 1991. Since the association between multilocus heterozygosity and fitness is quantitatively weak, it may be difficult to detect these effects in organisms where obtaining very large samples is not possible, due to the limited size of the population under study (Coltman and Slate, 2003). Similarly, the number of

individuals in any inbreeding class is small, reducing statistical power in comparisons involving *f*.

### Genomic aspects of HFCs

The genomic patterns responsible for generating HFCs are not well understood. One useful framework for understanding the genomic forces responsible for generating HFC's involves the local and general effects hypotheses (David *et al*, 1995). The local effects hypothesis suggests that HFC's are the result of linkage between marker and specific heterotic coding loci, whereas general effects suggests that total autozygosity (rather than heterozygosity at specific loci) predicts fitness. Under the general effects hypothesis, no correlation between *H* and fitness components should be detectable among identically inbred individuals because *f* is an index of selectable autozygosity (David, 1998; Hansson and Westerberg, 2002). It has been suggested that data sets that allow calculation of both inbreeding coefficients and individual heterozygosities may be helpful in distinguishing among hypotheses (Hansson and Westerberg, 2002).

In evaluating evidence for the local *versus* general effects hypothesis, we found no relationship between *H* and lifespan or recruitment within inbreeding classes of *G. scandens* or *G. fortis*. There was some evidence that recruits had higher heterozygosity at one level of inbreeding in *G. fortis*. Nevertheless, *H* is consistently larger in recruits than nonrecruits for the remaining comparisons, and all the correlation coefficients are positive. On the surface therefore, this appears to support the local effects hypothesis in that some of the marker loci may be linked to regions under selection, and larger sample sizes may have provided more statistical power to detect this. This reasoning assumes that *f* is a precise estimate of selectable autozygosity. However, the simulation results suggest that individuals within an inbreeding class can vary widely in autozygosity even when the number of linkage groups is large, as is thought to be the case in birds (Thorneycroft, 1975; Burt, 2002). Therefore, some of the variation in fitness due to heterosis that is not accounted for by *f* may be reflected in *H*. If so, then SSR HFC's within inbreeding classes may be due to variation among identically inbred individuals in true autozygosity, rather than a result of local effects.

With this data set, it is not possible to distinguish between the local and general effects hypotheses without a more complete understanding of the variance in autozygosity for a measured level of inbreeding, and this depends on an understanding of the size of linkage groups (Franklin, 1977). Although physical linkage groups in birds are likely to be small and numerous (Burt, 2002), bottlenecks and population structure can increase the level of linkage disequilibrium (Hansson and Westerberg, 2002), and perhaps falsely give the impression of local effects because each marker locus would represent a substantial portion of the genome. Thus, the general effects hypothesis can be rejected only after the variance in individual autozygosity for a known level of inbreeding has been accounted for (cf. Hansson and Westerberg, 2002, Table 2), or when linkage between specific marker loci and fitness effects can be established.

A major challenge when working with microsatellite or other highly polymorphic markers is to differentiate between statistical and biological significance (Hedrick, 1999). David (1998) has noted that HFC's are often very weak, and has questioned their biological importance. Although the overall correlation between lifespan and *H* is quantitatively weak in *G. fortis* and nonexistent in *G. scandens*, the details of our analyses suggest that both *H* and *f* are biologically meaningful even if they are imprecise estimators of global autozygosity, and ultimately, of fitness. The HFC observed in *G. fortis* explains much early mortality, and recruits tend to be more heterozygous than nonrecruits. Individuals with extreme levels of heterozygosity experience the greatest impact on lifespan. These results support the view that when adequate samples are available, multilocus heterozygosity data can provide an estimate of the level of selectable genetic diversity within a population, and can provide a useful complement to or substitute for data on inbreeding depression obtained through pedigree analysis.

## Acknowledgements

We thank the Galápagos National Parks Service and the Charles Darwin Research Station for their ongoing support, and I Lovette and G C Keyes for field assistance in 1991. This work was funded by the National Science Foundation (to PRG and BRG), the NSF/Sloan Foundation (to KP). We express our appreciation to J Watson and the Jin Laboratory at the University of Cincinnati for assistance with genotyping, and M Arnegard, J Gee, M Polak, and C Wilmhoff for useful comments and advice. We thank two anonymous reviewers for their input on a previous version of manuscript.

## References

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

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

Booth CL, Woodruff DS, Gould SJ (1990). Lack of significant associations between allozyme heterozygosity and phenotypic traits in the land snail *Cerion*. *Evolution* **44**: 210–213.

Britten HB (1996). Meta-analysis of the association between multilocus heterozygosity and fitness. *Evolution* **50**: 2158–2164.

Burt DW (2002). Origin and evolution of avian microchromosomes. *Cytogenet Genome Res* **96**: 97–112.

Coltman DW, Slate J (2003). Microsatellite measures of inbreeding: a meta-analysis. *Evolution* **57**: 971–983.

Coulson TN, Pemberton JP, Albon SD, Beaumont MA, Marshall TC, Slate J *et al* (1998). Microsatellites measure inbreeding depression and heterosis in red deer. *Proc R Soc London Ser B* **265**: 489–495.

David P (1998). Heterozygosity-fitness correlations: new perspectives on old problems. *Heredity* **80**: 531–537.

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

David P, Jarne P (1997). Context dependent survival differences among electrophoretic genotypes in natural populations of the marine bivalve *Spisula ovalis*. *Genetics* 335–344.

Frankham R (1998). Inbreeding and extinction: island populations. *Cons Biol* **12**: 665–675.

Frankham R, Lees K, Montgomery ME, England PR, Lowe EH, Briscoe DA (1999). Do population size bottlenecks reduce evolutionary potential? *Anim Cons* **2**: 255–260.
Franklin IR (1977). The distribution of the proportion of the genome which is homozygous by descent in inbred individuals. *Theoret Popul Biol* **11**: 60–80.
Goudet J, Keller L (2002). The correlation between inbreeding and fitness: does allele size matter? *Trends Ecol Evol* **17**: 201–202.
Grant PR, Grant BR (1992). Hybridization of bird species. *Science* **256**: 193–197.
Grant PR, Grant BR (2000). Non-random fitness variation in two populations of Darwin's finches. *Proc R Soc London B Biol Sci* **267**: 131–138.
Grant PR, Grant BR (2002). Unpredictable evolution in a 30 year study of Darwin's finches. *Science* **296**: 707–711.
Grant PR, Grant BR, Keller LF, Markert JA, Petren K (2003). Inbreeding and interbreeding in Darwin's finches. *Evolution* **57**.
Grant PR, Grant BR, Keller LF, Petren K (2000). Effects of El Nino events on Darwin's finch productivity. *Ecology* **81**: 2442–2457.
Grant PR, Grant BR, Petren K (2001). A population founded by a single pair of individuals: establishment, expansion, and evolution. *Genetica* **112**: 359–382.
Hansson B, Bensch S, Hasselquist D, Åkesson M (2001). Microsatellite diversity predicts recruitment of sibling great reed warblers. *Proc R Soc London B Biol Sci* **268**: 1287–1291.
Hansson B, Westerberg L (2002). On the correlations between heterozygosity and fitness in natural populations. *Mol Ecol* **11**: 2467–2474.
Hartl DL, Clark AG (1989). *Principles of Population Genetics*, 2nd edn. Sinauer Associates, Inc.: Sunderland, Massachusetts.
Hedrick P (1999). Perspective: highly variable loci and their interpretation in evolution and conservation. *Evolution* **53**: 313–318.
Hedrick P (2001). Conservation genetics: where are we now. *Trends Ecol Evol* **16**: 629–636.
Hedrick PW, Kalinowski ST (2000). Inbreeding depression in conservation biology. *Annu Rev Ecol Syst* **31**: 139–162.
Hoenig JM, Heisey DM (2001). The abuse of power: the pervasive fallacy of power calculations for data analysis. *Am Stat* **55**: 19–24.
Houle D (1994). Adaptive distance and the genetic basis of heterosis. *Evolution* **48**: 1410–1417.
Keller LF, Grant PR, Grant BR, Petren K (2001). Heritability of morphological traits in Darwin's finches: misidentified paternity and maternal effects. *Heredity* **87**: 325–336.
Keller LF, Grant PR, Grant BR, Petren K (2002). Environmental conditions affect the magnitude of inbreeding depression in survival of Darwin's finches. *Evolution* **56**: 1229–1239.
Keller LF, Waller DM (2002). Inbreeding effects in wild populations. *Trends Ecol Evol* **17**: 230–241.
Knaepkens G, Knapen D, Bervoets L, Hänfling B, Verheyen E, Eens M (2002). Genetic diversity and condition factor: a significant relationship in Flemish but not in German populations of the European bullhead (*Cottus gobio* L.). *Heredity* **89**: 280–287.
Lande R (1995). Mutation and conservation. *Cons Biol* **9**: 782–791.
Launey S, Hedgecock D (2001). High genetic load in the Pacific oyster *Crassostrea gigas*. *Genetics* **159**: 255–265.
Myrand B, Tremblay R, Sévigny J-M (2002). Selection against blue mussels (*Mytilus edulis* L.) homozygotes under various stressful conditions. *J Hered* **93**: 238–248.
Nei M (1978). Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics* **89**: 583–590.
Petren K (1998). Microsatellite primers from *Geospiza fortis* and cross-species amplification in Darwin's finches. *Mol Ecol* **7**: 1782–1784.
Petren K, Grant BR, Grant PR (1999a). Low extrapair paternity in the cactus finch (*Geospiza scandens*). *Auk* **116**: 252–256.
Petren K, Grant BR, Grant PR (1999b). A phylogeny of Darwin's finches based on microsatellite DNA length variation. *Proc Royal Soc London* **266**: 321–329.
Rowe G, Beebee TJC (2001). Fitness and microsatellite diversity estimates were not correlated in two outbred anuran poopulations. *Heredity* **87**: 558–565.
Savolainen O, Hedrick PW (1995). Heterozygosity and fitness: no association in scots pine. *Genetics* **140**: 755–766.
Thorneycroft HB (1975). A cytogenetic study of the white-throated sparrow *Zonotrichia albicollis*. *Evolution* **29**: 611–621.
Tsitrone A, Rousset F, David P (2001). Heterosis, marker mutational processes, and population inbreeding history. *Genetics* **159**: 1845–1859.
Vila C, Sundqvist A-K, Flagstad O, Seddon J, Björnerfeldt S, Kojola I *et al* (2002). Rescue of a severely bottlenecked wolf (*Canis lupus*) population by a single immigrant. *Proc R Soc London Ser B* **270**: 91–97.
Vucetich JA, Waite TA (1999). Erosion of heterozygosity in fluctuating populations. *Cons Biol* **13**: 860–868.
Westemeier RL, Brawn JD, Simpson SA, Esker TL, Jansen RW, Walk JW *et al* (1998). Tracking the long-term decline and recovery of an isolated population. *Science* **282**: 1695–1698.
Zouros E, Foltz DW (1987). The use of allelic isozyme variation for the study of heterosis. *Isozymes: Curr Topics Biolog Med Res* **13**: 1–59.

Copyright of Heredity is the property of Nature Publishing Company and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.