Correlations between heterozygosity and measures of genetic similarity: implications for understanding mate choice

S. Craig Roberts[1]*, Marie L. Hale[2] and Marion Petrie[3]

[1] *School of Biological Sciences, University of Liverpool, UK*

[2]*School of Biological Sciences, University of Canterbury, New Zealand*

[3]*School of Clinical Medical Sciences, University of Newcastle, UK*

*For contact: School of Biological Sciences, University of Liverpool, UK

E-mail: craig.roberts@liv.ac.uk   Tel: +44 151 414 7814   Fax: +44 151 414 4414

Running title: Heterozygosity, similarity and mate choice

Keywords: sexual selection, disassortative, directional selection, MHC, HLA

Abstract

There is currently considerable interest in testing the effects of genetic compatibility and heterozygosity on animal mate preferences. Evidence for either effect is rapidly accumulating, although results are not always clear-cut. However, correlations between mating preferences and either genetic similarity or heterozygosity are usually tested independently, and the possibility that similarity and heterozygosity may be confounded has rarely been taken into account. Here we show that measures of genetic similarity (allele-sharing, relatedness) may be correlated with heterozygosity, using data from 441 human individuals genotyped at major loci in the major histocompatibility complex, and 281 peafowl (*Pavo cristatus*) individuals genotyped at 13 microsatellite loci. We show that average levels of allele-sharing and relatedness are each significantly associated with heterozygosity in both humans and peafowl, that these relationships are influenced by the level of polymorphism, and that these similarity measures may correlate with heterozygosity in qualitatively different ways. We discuss the implications of these inter-relationships for interpretation of mate choice studies. It has recently become apparent that mating preferences for 'good genes' and 'compatible genes' may introduce discordant choice amongst individuals, since the optimal mate for one trait may not be optimal for the other, and our results are consistent with this idea. The inter-relationship between these measures of genetic quality also carries implications for the way in which mate choice studies are designed and interpreted, and generates predictions that can be tested in future research.

INTRODUCTION

Facilitated by the advent of molecular techniques, there has been a burgeoning interest in recent years in genetic influences on animal and human mate choice. Females, in particular, have been shown to choose between potential mates and to gain fitness benefits by selecting males who are superior to their competitors along some dimension of genetic quality (reviews in e.g. Bateson, 1983; Andersson, 1994; Jennions & Petrie, 2000; Tregenza & Wedell, 2000). While the exact nature of these good genes often remains unknown, three main categories of genetic influence on mate choice can be distinguished which have attracted considerable research interest (Mays & Hill, 2004). First, mating preferences may often be influenced by the expression of condition-dependent signals of quality, usually in males (e.g. Hasselquist *et al.,* 1996; Petrie *et al.,* 1991; Petrie, 1994; Gosling & Roberts, 2001). Females may increase their own reproductive success by using these traits to choose mates with "good genes" (Trivers, 1972) because offspring from these matings grow faster and are more likely to survive (Petrie, 1994), and may be more attractive in adulthood (Drickamer, 1992; Norris, 1993).

A second category of genetic influence is genetic compatibility, often described as the degree of genetic dissimilarity between potential mates (Zeh & Zeh, 1996; Tregenza & Wedell, 2000). In contrast with 'good genes', which can be seen as increasing fitness independently of the rest of the genome and show additive genetic variance, compatible genes increase fitness in the context of specific genotypes and show non-additive genetic variance (Puurtinen *et al.,* 2005). Disassortative mate choice may serve to avoid inbreeding (Pusey & Wolf, 1996) and increase heterozygosity in offspring, with consequent indirect fitness benefits to parents (Brown 1997, 1999; Garant *et al.,* 2005). Such effects are thought to be central to the evolution of polyandry (Zeh & Zeh, 1996, 2003; Jennions & Petrie, 2000; Tregenza & Wedell, 2000). One major source of genetic variation which has been implicated in vertebrate mate choice is at the major histocompatibility complex or MHC (reviews in Potts & Wakeland, 1993; Jordan & Bruford, 1998; Penn & Potts, 1999; Penn, 2002; Bernatchez & Landry, 2003). MHC-disassortative mating preferences have been shown in many taxa, including mice (e.g. Yamazaki *et al.,* 1976; Potts *et al.,* 1991; Roberts & Gosling, 2003), humans (Wedekind *et al.,* 1995; Wedekind & Furi, 1997), lizards (Olsson *et al.,* 2003), birds (Freeman-Gallant *et al.,* 2003) and fish (Landry *et al.,* 2001).

Thirdly, there is the possibility that individual heterozygosity may in itself be beneficial in terms of gaining matings (Brown, 1997). Compared with compatibility effects, where genetic quality in offspring depends on both parental genotypes, heterozygosity is an absolute trait on which females may tend to converge in their choice of suitable mates. Because of this, and in view of correlations between heterozygosity and condition-dependent

phenotypic traits (e.g. Ditchkoff *et al.,* 2001; Foerster *et al.,* 2003; Roberts *et al.,* 2005), heterozygosity preferences may be subsumed as a specific case of a "good genes" indicator trait, but we maintain Mays & Hill's (2004) distinction as it highlights particular genetic qualities and is currently the focus of increasing research interest. Heterozygote mating advantages may accrue through superior competitive ability or mate choice (Brown, 1997), although in the latter they are more likely to confer direct fitness benefits as males cannot pass their heterozygous condition to their offspring (Brown, 1997; Mays & Hill, 2004; but see Mitton *et al.,* 1993). In birds, for example, high individual heterozygosity is correlated with male territory size, song structure and seasonal reproductive success (Seddon *et al.,* 2004), and with fledging success (Foerster *et al.,* 2003), although in some cases intermediate heterozygosity is beneficial (Aparicio *et al.,* 2001). As with compatibility effects, the MHC has been a focus of much recent research (Brown, 1999), and MHC-heterozygote mating advantages have been found in several mammalian studies (e.g. Sauermann *et al.,* 2001; Ditchkoff *et al.,* 2001; Thornhill *et al.,* 2003; Roberts *et al.,* 2005).

Previous studies that have looked for genetic effects on mate choice have tended to focus on one of these three kinds of effect in isolation. However, it has recently become apparent that interactions between these effects may be a common and complicating factor in mate choice (Mays & Hill, 2004; Neff & Pitcher, 2005). In particular, the question of how individuals integrate information about 'good-genes' and compatibility in mating decisions has been modelled (Colegrave *et al.,* 2002) and examined experimentally (Roberts & Gosling, 2003). The strength of preference for genetic quality (additive genetic variation) or genetic distance (non-additive) may be influenced by the variability in either trait amongst potential mates and may contribute to the maintenance of variation in both, even under unitary directional selection (Roberts & Gosling, 2003). A similar interaction between compatibility and heterozygosity effects on mate preferences is likely to exist, but to date this has not been dealt with explicitly. Here we argue that a fuller understanding of their influence on mate choice may require that both effects are taken into account in any analysis, because they may often be inter-correlated.

Indeed, used as a coefficient of inbreeding, heterozygosity indicates levels of genetic diversity in a population and can be used to estimate genetic distance between populations (e.g. Chakraborty, 1984; Guerreiro *et al.,* 1994). Correlations may also hold at the individual level (Nei, 1978) and can be a consideration in calculating match probabilities in applied genetics, such as in forensic genetics (Ayres, 2000; Presciuttini *et al.,* 2002) and surveying the pool of suitable donors for tissue transplantation (Hata *et al.,* 1998; Marsh *et al.,* 2000). Despite this, the relationships between heterozygosity and measures of genetic similarity, and their consequences, have received little explicit attention in mate choice studies and we believe that they may be unappreciated by many working in this field.

In this paper, we examine correlations between heterozygosity and two measures of genetic similarity: the number of shared alleles and relatedness (*sensu* Queller & Goodnight, 1989). Numbers of shared alleles is the measure used in studies of human MHC-correlated mate preferences (Wedekind *et al.,* 1995; Wedekind & Furi, 1997; Jacob *et al.,* 2002; Thornhill *et al.,* 2003; Roberts *et al.,* 2005) and is qualitatively similar to band-sharing coefficients used in other vertebrate studies (e.g. Blomqvist *et al.,* 2002). Relatedness refines these estimates by controlling for allele frequencies, and is commonly used by behavioural and molecular ecologists in studies of non-random mating patterns. We first show how the correlation between heterozygosity and allele-sharing arises using an example of four diallelic loci, with alleles occurring at a range of frequencies from low polymorphism ($p = 0.9$, $q = 0.1$) to relatively high polymorphism ($p = 0.5$, $q = 0.5$). We then show that the relationship persists in more complex situations by examining heterozygosity and allele-sharing at the MHC in humans, and at a number of microsatellite loci in peafowl. The MHC is a pertinent example both because it is a primary focus of research on genetic influences on mating preferences and because it exhibits two properties that may confound a simple relationship between heterozygosity and genetic distance: a high degree of linkage disequilibrium and large heterogeneity in allele frequencies within populations (Marsh *et al.,* 2000). In view of the latter, we also examine the influence of heterozygosity at defined loci on relatedness (*sensu* Queller & Goodnight, 1989), which controls for allele frequency. Microsatellite data are the most commonly collected, selectively neutral data used by researchers to determine parentage and kinship within small groups and to determine relatedness at a population scale. Thus with these two data sets we can examine relationships between heterozygosity and genetic similarity both at loci under selection (MHC) and at selectively neutral loci (microsatellites). Finally, we discuss how inter-relationships between these measures may influence the outcome of studies investigating mate choice. Incorporating both effects in the interpretation of results is important to avoid the potential for arriving at incomplete conclusions and because it may generate new predictions that can be further tested in future studies.

METHODS

Human MHC genotyping:
We genotyped 441 individuals. To the best of our knowledge, none of these individuals were related: we specifically excluded a number of siblings from a larger sample that was available to us by randomly selecting one from each sib pair. Approximately 5ml of blood was collected using vacuettes (Greiner Bio-One Ltd, Stonehouse, Gloucestershire, UK) lined

with EDTA to prevent clotting. Genomic DNA was typed at three key MHC loci (HLA-*A, -B, -DRB1*), by polymerase-chain reaction using sequence-specific primers (PCR-SSP). Typing was done at the UK National Blood Service regional tissue typing laboratory in Newcastle. We used these three loci because, with over 300 known alleles each, they are the most polymorphic HLA loci according to the HLA database (www.ebi.ac.uk/imgt/hla/stats.html), and because they have been used in previous studies of HLA-associated mate preferences (Wedekind *et al.,* 1995; Wedekind & Furi, 1997; Jacob *et al.,* 2002; Thornhill *et al.,* 2003; Roberts *et al.,* 2005).

Of the 441 individuals (171 males, 270 females), 314 were heterozygous at all 3 loci, 105 were homozygous at one locus, 21 were homozygous at two loci and 1 was homozygous at all 3 loci. (Note that, because this latter individual was unique in our sample, his relatedness and allele matches data are plotted against heterozygosity in Figure 2 in order to illustrate all available data, but not included in the analyses of variance that we report). In comparing effects of heterozygosity on allele sharing and relatedness, we used an estimate of heterozygosity based on the number of loci at which an individual is heterozygous (possible range 0-3), following previous studies (Thornhill *et al.,* 2003; Roberts *et al.,* 2005). We recorded 19 different alleles at HLA-*A*, 36 alleles at –*B* and 18 at –*DRB1*, and heterozygosity values were 0.87, 0.92 and 0.87, respectively.

For each individual, we calculated the number of shared alleles with each of the other individuals using SHAREDST (available from www2.biology.ualberta.ca/jbrzusto/sharedst.php), and then averaged these pair-wise values to obtain a mean score. A high score indicates that the individual shares a relatively large number of alleles, on average, with other members of the population. In common with previous studies of HLA-based mate preferences (Wedekind *et al.,* 1995; Wedekind & Furi, 1997; Jacob *et al.,* 2002; Thornhill *et al.,* 2003; Roberts *et al.,* 2005), two individuals who are both homozygous for the same allele at a given locus were scored as sharing 2 alleles for that locus. Average allele sharing across all individuals was 1.21 (SD=0.31, range 0.31-2.0), which is similar to that recorded in their sample by Thornhill *et al.,* (2003) (mean=1.25; the other studies do not provide this information).

We also calculated pair-wise genetic relatedness scores according to Queller & Goodnight's method (Queller & Goodnight, 1989), using the program RELATEDNESS (version 5.0.8; available from www.gsoftnet.us/GSoft.html). The relatedness calculation is given by:

$$R = \frac{\sum_{x} \sum_{k} \sum_{l} \left(P_y - P^*\right)}{\sum_{x} \sum_{k} \sum_{l} \left(P_x - P^*\right)}$$

where *x* indexes individuals in the data set, *k* indexes loci and *l* indexes allelic position (i.e. *l* =1 or 2 for a diploid individual, 1 only for a haploid), and where P_x is the frequency within the *x*th individual of the allele found at *x's* locus *k* and allelic position *l*, P_y is the frequency of that same allele in the set of "partners" of *x* (i.e. individuals being compared with the *x*th individual), and P* is the frequency of the allele in the population at large (with all putative relatives of *x* excluded). As we did for allele sharing, we averaged pair-wise relatedness scores with all other individuals, to calculate a mean relatedness score for each individual. This measure of relatedness takes into account allelic frequencies as well as allele sharing, giving more weight to sharing of rare alleles than sharing of relatively common alleles. Positive scores indicate individuals are more similar, on average, to other members of the population than individuals with negative scores.

We did not use the bias correction function for exclusion of each individual (and its putative relatives, if any) from the calculation of population allele frequencies, as all our individuals were unrelated and the sample size was large. This latter point is important both because the need for bias correction is reduced in larger samples as the contribution of each individual becomes negligible (Queller & Goodnight, 1989) and because our sample size exceeded the capacity of the program (maximum 127 groups, where we have 441 family groups with one member in each). However, to check that correlations between heterozygosity and mean pair-wise relatedness were not influenced or determined by this potential bias, we re-ran the analysis, this time using the bias-correction facility but replacing a family identifier with subject's sex as the grouping variable (i.e. when current individual X was male, allele frequencies were calculated from females only). In this re-analysis, the same negative relationship between heterozygosity and relatedness was found, indicating that the reported effect could not be due to this bias.

We analysed variation in average allele sharing and relatedness scores in relation to average heterozygosity using one-way analysis of variance. Data were approximately normally distributed (Kolmogorov-Smirnov tests, both *P*>0.2). We compared the regressions between mean allele sharing and relatedness scores for individuals of variable heterozygosity by a one-way analysis of covariance (ANCOVA, Sokal & Rohlf, 1995) with mean relatedness scores as the dependent variable, mean allele sharing as the covariate and number of heterozygous loci as the factor. We first tested for heterogeneity amongst slopes, and since these were not significantly different, proceeded to test for differences amongst intercepts using analysis of covariance. Analyses were carried out using SPSS version 12 and all tests are non-directional.

Peafowl microsatellite genotyping:

We genotyped 46 adult male peafowl and 235 adult females at 13 microsatellite loci: PC3, PC41, PC46, PC142, PC151, PC125, PC281, PC256, PC9, PC148, PC36, PC159 and PC243 (Hale *et al.,* 2004). DNA extraction, primer details, PCR amplification conditions, fragment detection and analysis were as described in Hale *et al.,* (2004). The 13 microsatellite loci ranged in polymorphism level from very low (H_E = 0.05) to moderately polymorphic (H_E = 0.70), and varied between two and six alleles per locus. None of the loci for which we developed primers were highly polymorphic, probably a result of the fact that our population of semi-captive, non-native animals would have descended from a small founder population. Each male was mated with eight to ten females, and each female was mated with one or two males, as part of a larger experiment to examine paternal and maternal genetic effects (including heterozygosity and parental genetic similarity) on offspring size, growth and sex ratio in peafowl. The number of shared alleles and Queller & Goodnight's symmetrical relatedness between parents were calculated as above for each of the matings that actually took place, a total of 388 matings. Individual heterozygosity ranged from heterozygous at zero loci to heterozygous at nine of the 13 loci.

To examine the effects of polymorphism on the relationship between heterozygosity and genetic similarity, we split the data set into two subsets: 1) the four most polymorphic loci only (PC142, PC151, PC256 and PC159; H_E range = 0.49 to 0.70, average H_E = 0.56), and 2) the four least polymorphic loci only (PC3, PC9, PC36 and PC281; H_E range = 0.05 to 0.20, average H_E = 0.15). For each of these subsets we examined the relationship between individual heterozygosity (the number of heterozygous loci out of the four loci examined) and genetic similarity using one-way analysis of variance.

RESULTS

Correlation between heterozygosity and allele-sharing

At its simplest, the probability of individuals sharing the same genotype at a single diallelic locus where *p*=*q*=0.5 are p^2 for homozygotes and 2*pq* for heterozygotes. Thus, the probability of two individuals sharing the same genotype is 0.25 for homozygotes and 0.5 for heterozygotes.

A similar trend is evident by calculating the number of matching alleles between members of a population. If we extend this to a diallelic locus in Hardy-Weinberg equilibrium

where the allele frequencies are not equal, then the average number of shared alleles for the dominant homozygote will be $2p^2 + 2pq$ (i.e. two shared alleles with the dominant homozygote which occurs at a frequency of p^2 and one shared allele with heterozygotes which occur at a frequency of $2pq$, no shared alleles with the recessive homozygote). Following the same reasoning, the number of shared alleles for the recessive homozygote will be $2pq + 2q^2$, and the number of shared alleles for heterozygotes will be $p^2 + 4pq + q^2$. Thus the combined number of shared alleles for all homozygotes (a weighted average across both homozygous genotypes based on each homozygote's genotype frequency) will be $(2p^4 + 2p^3q + 2pq^3 + 2q^4) / (p^2 + q^2)$. From these equations we can see the effect that polymorphism (defined as expected heterozygosity (H_E) under Hardy-Weinberg equilibrium) has on the relationship between heterozygosity and allele-sharing. At the highest level of polymorphism possible at a diallelic locus ($p = q = 0.5$, $H_E = 0.50$), the average number of shared alleles will be 1.00 for homozygotes, and 1.50 for heterozygotes. In contrast, at a relatively low level of polymorphism ($p = 0.9$, $q = 0.1$, $H_E = 0.18$), the average number of shared alleles will be 1.78 for homozygotes and only 1.18 for heterozygotes. Thus the relationship between heterozygosity and allele-sharing is positive at high levels of polymorphism and negative at loci with low polymorphism and is caused by the interaction between the frequency of homozygotes for particular alleles and the frequencies of these alleles. For example, at very uneven allele frequencies, homozygotes for the common allele will be plentiful, resulting in higher average allele-sharing with other population members than when allele frequencies are equal. In contrast, uneven allele frequencies will mean that frequency of highly heterozygous individuals will be low, resulting in lower allele-sharing than when allele frequencies are more even. Figure 1 illustrates this same point for a theoretical case with 4 diallelic loci, all with the same allele frequencies.

Relationships at the human Major Histocompatibility Complex
Our results using the highly polymorphic human MHC data demonstrate a strong positive relationship between heterozygosity and allele-sharing as predicted by Fig. 1, even at polyallelic loci where there is heterogeneity in allele frequency and significant linkage disequilibria. Allele sharing with other members of the population varies significantly with degree of individual heterozygosity ($F_{2,437}=8.31$, P<0.001, $\eta^2=0.037$), being higher in relatively heterozygous individuals than in homozygotes (Fig.2a). Individual heterozygosity is also associated with average relatedness ($F_{2,437}=12.93$, P<0.001; $\eta^2=0.056$, Fig.2b), although here the relationship is negative. Although the effect sizes are small, the main point is that the two measures are related to heterozygosity in different directions. Including the single individual who was found to be homozygous at all three loci does not alter these

effects ($F_{3,437}$=5.79, P<0.001, η^2=0.038 and $F_{3,437}$=8.88, P<0.001, η^2=0.057, for allele sharing and relatedness, respectively).

It was surprising to find that allele sharing and relatedness are related to heterozygosity in opposite directions, given that the two measures of genetic similarity are inter-related. Indeed, in our sample they are positively correlated (Pearson r=0.89, n=441, P<0.001). Figure 3 plots this correlation between allele-sharing and relatedness in the same dataset, and shows that relatedness decreases with heterozygosity for a given degree of allele-sharing. Analysis of covariance shows that, although the slopes are equivalent for the three groups ($F_{2,434}$=1.18, P=0.31; partial η^2=0.005), the intercepts are significantly different ($F_{2,436}$=1699.9, P<0.001; partial η^2=0.89). These different elevations are responsible for generating the differences in the mean values for allele-sharing and relatedness, for each heterozygosity category, observed in Figure 2. The explanation for why relatedness values alter with the degree of heterozygosity (i.e. have different elevations) appears to be due to the nature of the algorithm of the RELATEDNESS program; we will return to this point later.

An alternative to our definition of heterozygosity as the number of loci for which each individual is heterozygous is Coltman *et al.*'s standardised heterozygosity measure, H_s (Coltman *et al.,* 1999), which weights heterozygosity at each locus according to the average heterozygosity at that locus. Although this has some advantages over simply summing the number of heterozygous loci, it is most useful when individuals are not all typed at the same loci (Coltman *et al.,* 1999; Seddon *et al.,* 2004). Because this is not the case here, and because we use a relatively small number of loci, calculation of H_s scores yields only a small additional amount of variation between individuals compared to using the number of heterozygous loci. Despite this, we find that H_s correlates positively with average allele sharing (Spearman's rank correlation, r_s= 0.191, n=441, P<0.001) and negatively with average relatedness (r_s= -0.236, P<0.001). These results are consistent with our main analyses.

Relationships at peafowl microsatellite loci

As with the human data, there is a negative relationship across all 13 microsatellite loci between individual heterozygosity and relatedness (Spearman's rank correlation, r_s = -0.187, n = 388, P<0.001). However there was no relationship between individual heterozygosity and the number of shared alleles (r_s = -0.062, n = 388, P=0.222). The reason for this lack of relationship becomes clear when we examine the relationship between heterozygosity and allele-sharing in the two subsets of loci: the four most polymorphic loci and the four least polymorphic loci. The relationship between heterozygosity and the number of shared alleles is strongly negative in the low polymorphic subset ($F_{3,383}$ = 49.79, P< 0.001, η^2 = 0.280), but appears weakly positive in the moderately polymorphic subset ($F_{3,383}$ = 1.71, P=0.14, Fig. 4a,

$\eta^2 = 0.018$), as predicted by Figure 1. Thus when data are averaged over all loci with varying levels of polymorphism, the negative relationship between the low polymorphic loci is masked by the weakly positive relationship at the moderately polymorphic loci, so that overall no relationship is detected.

The same situation can be seen if we re-examine the relatedness effect, this time taking into account variable levels of polymorphism. When analysing only the four least polymorphic loci, the relationship between heterozygosity and relatedness is strongly negative ($F_{3,383}$ = 45.34, P<0.001, η^2 = 0.262), but the relationship disappears amongst only the four most polymorphic loci ($F_{3,383}$ = 0.65, P=0.63, η^2 = 0.007, Fig. 4b).

The relationship between relatedness and heterozygosity

Allele sharing and relatedness are related to heterozygosity in opposite directions (e.g. Fig,2), despite the two measures of genetic similarity being inter-related (Fig.3). How does this come about? The explanation appears to be due to the fact that heterozygous individuals have more unique alleles than homozygous individuals, for a given number of shared alleles. Unique alleles (those present in the focus individual but not the partner) have a large impact on the computation of relatedness, adding a potentially large negative component (equal to the unique allele's frequency in the total population) to the numerator of the calculation (see methods). The result is that, for a given number of shared alleles, relatedness scores are lower for heterozygotes than homozygotes (see Fig.3). We expand on this below to demonstrate how this effect arises.

Heterozygosity has two effects on the calculation of relatedness that are absent in calculations of allele-sharing. The first, and minor, effect is that homozygous loci contribute twice as much to the denominator of the relatedness calculation as heterozygous loci (assuming equal allele frequencies). If allele frequencies are unequal, then homozygotes still contribute more but not necessarily twice as much. The end result is that individuals homozygous at many loci have a larger denominator and thus lower relatedness than heterozygous individuals when the number of shared alleles is held constant. However, there is also a second effect acting in the opposite direction that is much greater in magnitude. Heterozygous individuals have more unique alleles (alleles present in the focus individual but not the partner) than homozygous individuals when shared alleles are held constant (see example below). These unique alleles have a large impact on the relatedness calculation: each adds a negative component to the numerator which is equal to the unique allele's frequency in the total population.

This is most easily demonstrated using a simple example. Assume we have 3 loci (A, B, C), with 3 alleles at each locus (A1, A2, A3 etc) and allele frequencies at each locus are

equal (p = q = r = 0.33). Assume we have one potential partner and three focus individuals – one with one heterozygous locus (Het_1), one with 2 heterozygous loci (Het_2), and one with 3 heterozygous loci (Het_3). We will hold shared alleles constant, so for each individual, the number of shared alleles = 1. With these restrictions there is only one possible genotype for each individual, ignoring order of the loci (order is not important in this example as allele frequencies are the same at all loci).

Our genotypes are then (all combinations have 1 shared allele with partner, unique alleles are in bold):

Partner   A1,A2   B1,B2   C1,C1

Focus individuals:

Het_1   A1,A1   B1,B1   C1,**C2**

Het_2   A1,A1   B1,**B3**   C1,**C2**

Het_3   A1,**A3**   B1,**B3**   C1,**C2**

Notice that the number of unique alleles (those not present in partner) increases with increasing heterozygosity. This is not simply due to an arbitrary choice of genotypes: these are the only genotypes for these loci that will give a shared alleles value of 1 over all three loci (at each locus, the allele in the first position is the same in focus and partner, the allele in the second position is different); the only way to increase heterozygosity while keeping shared alleles constant is to introduce unique alleles.

The relatedness calculation (Queller & Goodnight, 1989) for each of these focus individuals to our one partner, then is (with some rounding error):

Het_1: $\dfrac{(0.5 - 0.33) + (0.5 - 0.33) + (0.5 - 0.33) + (0.5 - 0.33) + (1.0 - 0.33) + \mathbf{(0.0 - 0.33)}}{(1.0 - 0.33) + (1.0 - 0.33) + (1.0 - 0.33) + (1.0 - 0.33) + (0.5 - 0.33) + (0.5 - 0.33)}$

   = $\dfrac{1.01}{3.02}$ = 0.334

Het 2: $\dfrac{(0.5 - 0.33) + (0.5 - 0.33) + (0.5 - 0.33) + \mathbf{(0.0 - 0.33)} + (1.0 - 0.33) + \mathbf{(0.0 - 0.33)}}{(1.0 - 0.33) + (1.0 - 0.33) + (0.5 - 0.33) + (0.5 - 0.33) + (0.5 - 0.33) + (0.5 - 0.33)}$

   = $\dfrac{0.52}{2.02}$ = 0.257

Het 3: $\underline{(0.5 - 0.33) + \textbf{(0.0 – 0.33)} + (0.5 - 0.33) + \textbf{(0.0 – 0.33)} + (1.0 - 0.33) + \textbf{(0.0 – 0.33)}}$
$(0.5 - 0.33) + (0.5 - 0.33) + (0.5 - 0.33) + (0.5 - 0.33) + (0.5 - 0.33) + (0.5 - 0.33)$

$= \quad \dfrac{0.02}{1.02} \quad = \quad 0.019$

Notice that (i) the denominator decreases with increasing heterozygosity (due to homozygous loci effectively being double-counted because their frequency is 1.0 at each allelic position rather than 0.5), but also that (ii) this effect is overwhelmed by the decrease in the numerator caused by the negative components from the unique alleles (shown in bold). The end result is that relatedness decreases with increasing heterozygosity.

These calculations above are for asymmetrical relatedness (i.e. from the focus individuals' point of view only) but the relationship is the same for symmetrical relatedness – reversing focus and partner individuals still gives lower relatedness for the more heterozygous partners. Importantly, they also assume that allele frequencies are equal (that is, the highest level of polymorphism possible at these loci). However, changing the level of polymorphism has relatively little effect. Figure 5 plots the relationship between relatedness and heterozygosity for the above example, across a range of polymorphisms and shows that increasing heterozygosity always results in lower relatedness for a given number of shared alleles. These relationships are not linear, because unique alleles have relatively greater impact at lower allele frequencies, so that relatedness decreases sharply at very low levels of polymorphism for individuals with any heterozygous loci. The steepness of this drop-off increases with the proportion of heterozygous loci. The important point, however, is that relatedness is always lower for individuals with more heterozygous loci when shared alleles are kept constant, regardless of the level of polymorphism.

However, the relationship between heterozygosity and relatedness may be influenced by polymorphism where the number of shared alleles between pairs is variable and not held constant as above, such as in our human and peafowl data. The combined results of these datasets suggest that both very low polymorphism (some peafowl microsatellite loci) and very high polymorphism (human MHC) indeed result in a strong negative relationship between heterozygosity and relatedness. The exception in our dataset is in loci with moderate polymorphism, which show no such relationship. We believe the explanation for this is that individuals with high heterozygosity at loci with moderate polymorphism will possess fewer unique alleles than will heterozygotes at loci with extremely high or low polymorphism. When the level of polymorphism is very low, heterozygotes will be very rare, and are therefore unlikely to be mated with partners sharing both alleles. Thus, when loci have very low polymorphism, each additional locus at which an individual is

heterozygous will likely add another unique allele, and therefore another negative component, to the relatedness calculation (see methods). At very high polymorphism, heterozygotes are most likely to be mated to heterozygotes, but their partners will generally be heterozygous for other alleles. Therefore, at very high levels of polymorphism (such as human MHC) each locus at which an individual is heterozygous will also likely add unique alleles to the relatedness calculation. However, at intermediate levels of polymorphism (e.g. loci with only two or three equally common alleles), heterozygotes are likely to be paired with heterozygotes that share the same alleles, so the number of unique alleles will be low.

DISCUSSION

Multiple influences on mate choice

Overall, our results demonstrate the principle that heterozygosity and two of the most commonly used measures of genetic similarity may often be inter-correlated. This correlation clearly carries the potential to influence mate choice decisions at an individual level. For example, our results show that it is more likely that relatively heterozygous males will share alleles with females than homozygotes at highly polymorphic loci. If male heterozygosity is valued by females, the pool of genetically dissimilar mates will consequently be reduced and females would need to become more selective with respect to choosing genetically dissimilar males than in systems where selection for heterozygosity in mates is lower or neutral. On the other hand, if heterozygosity and low relatedness are valued by females, these two desirable qualities appear to be more congruent.

It is important to note that the inter-correlation between the two forms of genetic quality does not imply that it is impossible for individuals to maximise both aspects, since the correlations are based on averages across all pair-wise combinations. In other words, a female may theoretically find a heterozygote who shares no alleles with her, or indeed a homozygote who shares exactly the same set of alleles. However, it does mean that, in cases where females have limited opportunity for choice, they may be unable to find the ideal mate on both continua and may be forced to trade-off one aspect against the other.

As well as introducing complexity and variability in mating decisions between individual females, the inter-relationship between heterozygosity and genetic dissimilarity introduces a potential hidden cost to the apparent benefits gained by disassortative mating. While disassortative mating preferences result in more heterozygous offspring which may be healthier (Apanius *et al.,* 1997) and display attractive secondary sexual traits (e.g. Ditchkoff *et al.,* 2001; Foerster *et al.,* 2003; Roberts *et al.,* 2005), these offspring will also tend to share more alleles with the average opposite-sex individual and may thus have to be more

selective in their own mate choice. This kind of inter-generational effect is likely to be a general one in any system where the genes governing compatibility also have direct or pleiotropic fitness effects. This will include, but is not limited to, the MHC, where it is already known that natural selection for disease resistance may favour optimal levels of heterozygosity because of the degree of T-cell depletion during ontogenetic thymic selection (Nowak *et al.,* 1992; but see Borghans *et al.,* 2003). This led Penn & Potts (1999) to suggest that females should prefer to mate with males having intermediate, rather than maximal, levels of MHC-dissimilarity. Their suggestion has been supported by subsequent evidence from MHC-correlated odour preferences in humans (Jacob *et al.,* 2002) and three-spined sticklebacks, *Gasterosteus aculeatus* (Aeschlimann *et al.,* 2003; Milinski, 2003).

Interpreting patterns of mate choice

Our results raise a general methodological issue which is fundamental to the interpretation and design of mate choice studies. Since there is the potential for trade-offs between heterozygosity and measures of genetic similarity in mate choice decisions, both kinds of effect should be taken into account, wherever possible, in analyses of mate preferences and mating success. Although analysis of either trait in isolation will in many cases reveal interesting relationships, clear-cut or linear preferences for either trait may not always emerge.

Particularly in field studies, taking both traits into account will most usually be achieved by including each in statistical models to reveal interactions or independent effects (e.g. Blais *et al.,* 2004; Cordero *et al.,* 2004; Roberts *et al.,* 2005). In some cases, however, the nature of the analysis may completely remove effects of the other variable. One example of this can be found in analyses of the influence of MHC-dissimilarity on women's preferences for male body odour (Wedekind *et al.,* 1995). While a first comparison of scores given by normally-cycling women to MHC-similar and dissimilar male odours indicated a preference for the latter, this could potentially have been confounded by male heterozygosity (although it would perhaps indicate a preference for homozygous men). However, a second comparison using men as the unit of analysis effectively removes this possibility, since in this case all male effects are constant and the only variant is the relative number of alleles shared with the women making preference judgements (see also Roberts *et al*., *in press,* for the same analysis in facial preferences).

The finding that allele-sharing and relatedness are differentially correlated with heterozygosity, at least in our human dataset, carries a further consideration for interpretation of mate choice studies. While we have suggested that this is at least partly a computational effect resulting from the weighting given to unique alleles in the pairwise

calculation of relatedness, the question remains whether it may also be biologically meaningful. Researchers aiming to integrate the two kinds of genetic influence must therefore ideally distinguish what measure of genetic similarity females actually use. That is, heterozygosity aside, do they assess mate complementarity based on simple allele-sharing or do they also take into account the frequencies of the alleles involved?

Predictions

Finally, these inter-relationships between measures of mate quality also suggest several testable predictions. First, one might expect non-linearity in relationships between heterozygosity and male mating success, even if there is linearity between individual heterozygosity and secondary sexual traits. For example, this could potentially explain a previously-reported and unexplained quadratic effect of heterozygosity on mating success in spotless starlings, *Sturnus unicolor* (Aparicio *et al.,* 2001). Here, a negative relationship was reported between homozygosity and sexually-selected throat feathers, in accordance with the good-genes as heterozygosity hypothesis (Brown, 1997). In contrast, mating success was highest amongst males of intermediate heterozygosity level, and females were more likely to engage in extra-pair matings if their primary partner was extremely homozygous or heterozygous. We cannot say for certain whether this pattern was due to inter-correlations between heterozygosity and genetic similarity, but it seems a possible explanation and is supported by other studies which find links between the tendency to seek extra-pair copulations and high within-pair genetic similarity (Blomqvist *et al.,* 2002; Foerster *et al.,* 2003; Freeman-Gallant *et al.,* 2003; Eimes *et al.,* 2005) or, at a population level, low genetic variability (Petrie *et al.,* 1998). Aparicio *et al.* (2001) took this curvilinear relationship between heterozygosity and mating success as evidence against the good-genes as heterozygosity hypothesis. However, our results would suggest that their data are consistent with this hypothesis acting in conjunction with a preference for genetic dissimilarity.

Second, there may be condition-dependent effects in the strength of female preferences for either trait. For example, if male heterozygosity confers direct benefits in terms of offspring survival (e.g. in blue tits, *Parus caeruleus,* Foerster *et al.,* 2003), females might attach greater weight to male heterozygosity when resources are limited, or when the female is in poor condition. Seddon *et al.* (2004) found a positive correlation in subdesert mesites (*Monias benschi*) between male heterozygosity and both territory size and the number of surviving young at the end of the breeding season. It would be interesting to know whether and how such relationships vary across good and bad years.

Third, individual female heterozygosity should correlate with choosiness. Since heterozygous females will share more alleles, on average, with other males in the population, they would be expected to be more selective for genetic dissimilarity in mates

than relatively homozygous females. Thus, at least if the opportunity for choice is limited, female heterozygosity might be correlated with the degree of allele-sharing with their social partner and with the incidence of extra-pair matings. Some effects of female heterozygosity on reproductive success are known: it is correlated with larger clutch size in blue tits (Foerster *et al.,* 2003), and with hatching success in spotless starlings (Cordero *et al.,* 2004). In the latter, another quadratic effect was detected such that eggs laid by females with extremely high or low heterozygosity were less likely to hatch (see also Hansson, 2004), who found a curvilinear effect of relatedness on hatching success, and attributed decreasing hatching rates at high relatedness to genome-wide homozygosity). However, studies that have tested heterozygosity effects on mate preferences have so far tended to concentrate predominantly on males, and the heterozygosity-genetic similarity relationship indicates that greater regard to females needs to be integrated in the future.

Acknowledgments

References

Aeschlimann, P. B., Haberli, M. A., Reusch, T. B. H., Boehm, T. & Milinski, M. 2003. Female sticklebacks *Gasterosteus aculeatus* use self-reference to optimize MHC allele number during mate selection. *Behav.Ecol.Sociobiol.* **54**: 119-126.

Andersson, M. 1994. *Sexual Selection*. Princeton, New Jersey: Princeton University Press.

Apanius, V., Penn, D., Slev, P., Ruff, L. R. & Potts, W. K. 1997. The nature of selection on the major histocompatibility complex. *Crit.Rev.Immunol.* **17**: 179-224.

Aparicio, J. M., Cordero, P. J. & Veiga, J. P. 2001. A test of the hypothesis of mate choice based on heterozygosity in the spotless starling. *Anim.Behav.* **62**: 1001-1006.

Ayres, K. L. 2000. A two-locus forensic match probability for subdivided populations. *Genetica* **108**: 137-143.

Bateson, P. P. G. 1983, *Mate Choice*. Cambridge: Cambridge University Press.

Bernatchez, L. & Landry, C. 2003. MHC studies in nonmodel vertebrates: what have we learned about natural selection in 15 years? *J.Evol.Biol.* **16**: 363-377.

Blais, J., Rico, C. & Bernatchez, L. 2004. Nonlinear effects of female mate choice in wild threespine sticklebacks. *Evolution* **58**: 2498-2510.

Blomqvist, D., Andersson, M., Küpper, C., Cuthill, I. C., Kis, J., Lanctot, R. B., Sandercock, B. K., Székely, T., Wallander, K. & Kempenaers, B. 2002. Genetic similarity between mates and extra-pair parentage in three species of shorebirds. *Nature* **419**: 613-615.

Borghans, J. A. M., Noest, A. J. & Boer, R. J. D. 2003. Thymic selection does not limit the individual MHC diversity. *Eur.J.Immunol.* **33**: 3353-3358.

Brown, J. L. 1997. A theory of mate choice based on heterozygosity. *Behav.Ecol.* **8**: 60-65.

Brown, J. L. 1999. The new heterozygosity theory of mate choice and the MHC. *Genetica* **104**: 215-221.

Chakraborty, R. 1984. Relationship between heterozygosity and genetic distance in the three major races of man. *Amer.J.Phys.Anthropol.* **65**: 249-258.

Colegrave, N., Kotiaho, J. S. & Tomkins, J. L. 2002. Mate choice or polyandry: reconciling genetic compatibility and good genes sexual selection. *Evol.Ecol.Res.* **4**: 911-917.

Coltman, D. W., Pilkington, J. G., Smith, J. A. & Pemberton, J. M. 1999. Parasite-mediated selection against inbred Soay sheep in a free-living, island population. *Evolution* **53**: 1259-1267.

Cordero, P. J., Aparicio, J. M. & Veiga, J. P. 2004. Parental genetic characteristics and hatching success in the spotless starling, *Sturnus unicolor*. *Anim.Behav.* **67**: 637-642.

Ditchkoff, S. S., Lochmiller, R. L., Masters, R. E., Hoofer, S. R. & van den Bussche, R. A. 2001. Major-histocompatibility-complex-associated variation in secondary sexual traits of white-tailed deer (*Odocoileus virginianus*): Evidence for good-genes advertisement. *Evolution* **55**: 616-625.

Drickamer, L. C. 1992. Oestrous female house mice discriminate dominant from subordinate males and sons of dominant from sons of subordinate males by odour cues. *Animal Behaviour* **43**: 868-870.

Eimes, J. A., Parker, P. G., Brown, J. L. & Brown, E. R. 2005. Extrapair fertilization and genetic similarity of social mates in the Mexican jay. *Behav.Ecol.* **16**: 456-460.

Foerster, K., Delhey, K., Johnsen, A., Lifjeld, J. T. & Kempenaers, B. 2003. Females increase offspring heterozygosity and fitness through extra-pair matings. *Nature* **425**: 714-717.

Freeman-Gallant, C. R., Meguerdichian, M., Wheelwright, N. T. & Sollecito, S. V. 2003. Social pairing and female mating fidelity predicted by restriction fragment length polymorphism similarity at the major histocompatibility complex in a songbird. *Mol.Ecol.* **12**: 3077-3083.

Garant, D., Dodson, J. D. & Bernatchez, L. 2005. Offspring genetic diversity increases fitness of female Atlantic salmon (*Salmo salar*). *Behav.Ecol.Sociobiol.* **57**: 240-244.

Gosling, L. M. & Roberts, S. C. 2001. Scent-marking by male mammals: cheat-proof signals to competitors and mates. *Adv.Stud.Behav.* **30**: 169-217.

Guerreiro, J. F., Santos, E. J. M. D. & Santos, S. E. B. D. 1994. Effect of average heterozygosity on the genetic distance of several Indian tribes from the Amazon region. *Ann.Hum.Biol.* **21**: 589-595.

Hale, M. L., Petrie, M. & Wolff, K. 2004. Polymorphic microsatellite loci in peafowl (*Pavo cristatus*). *Mol.Ecol.Notes* **4**: 528-530.

Hansson, B. 2004. Marker-based relatedness predicts egg-hatching failure in great reed warblers. *Conservation Genetics* **5**: 339-348.

Hasselquist, D., Bensch, S. & von Schantz, T. 1996. Correlation between male song repertoire, extra-pair paternity and offspring survival in the great reed warbler. *Nature* **381**: 229-232.

Hata, Y., Cecka, J. M., Takemoto, S., Ozawa, M., Cho, Y. W. & Terasaki, P. I. 1998. Effects of changes in the criteria for nationally shared kidney transplants for HLA-matched patients. *Transplantation* **65**: 208-212.

Jacob, S., McClintock, M. K., Zelano, B. & Ober, C. 2002. Paternally inherited HLA alleles are associated with women's choice of male odor. *Nat.Genet.* **30**: 175-9.

Jennions, M. D. & Petrie, M. 2000. Why do females mate multiply? A review of the genetic benefits. *Biol.Rev.* **75**: 21-64.

Jordan, W. C. & Bruford, M. W. 1998. New perspectives on mate choice and the MHC. *Heredity* **81**: 239-245.

Landry, C., Garant, D., Duchesne, P. & Bernatchez, L. 2001. 'Good genes as heterozygosity': the major histocompatibility complex and mate choice in Atlantic salmon (*Salmo salar*). *Proc.R.Soc.Lond.B* **268**: 1279-1285.

Marsh, S. G. E., Parham, P. & Barber, L. D. 2000. *The HLA Facts Book*. London: Academic Press.

Mays, H. L. J. & Hill, G. E. 2004. Choosing mates: good genes versus genes that are a good fit. *Trend.Ecol.Evol.* **19**: 555-559.

Milinski, M. 2003. The function of mate choice in sticklebacks: optimizing MHC genetics. *J.Fish Biol.* **63 (Suppl.A)**: 1-16.

Mitton, J. B., Schuster, W. S. F., Cothran, E. G. & Fries, J. C. D. 1993. Correlation between the individual heterozygosity of parents and their offspring. *Heredity* **71**: 59-63.

Neff, B. D. & Pitcher, T. E. 2005. Genetic quality and sexual selection: an integrated framework for good genes and compatible genes. *Mol.Ecol.* **14**: 19-38.

Nei, M. 1978. Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics* **89**: 583-590.

Norris, K. 1993. Heritable variation in a plumage indicator of viability in male great tits, *Parus major*. *Nature* **362**: 537-539.

Nowak, M. A., Tarczy-Hornoch, K. & Austyn, J. M. 1992. The optimal number of major histocompatibility complex molecules in an individual. *Proc.Natl.Acad.Sci.* **89**: 10896-10899.

Olsson, M., Madsen, T., Nordby, J., Wapstra, E., Ujvari, B. & Wittsell, H. 2003. Major histocompatibility complex and mate choice in sand lizards. *Proc. R. Soc. Lond. B (Suppl.)* **270**: S254-S256.

Penn, D. J. 2002. The scent of genetic compatibility: sexual selection and the major histocompatibility complex. *Ethology* **108**: 1-21.

Penn, D. J. & Potts, W. K. 1999. The evolution of mating preferences and major histocompatibility complex genes. *American Naturalist* **153**: 145-164.

Petrie, M. 1994. Improved growth and survival of offspring of peacocks with more elaborate trains. *Nature* **371**: 598-599.

Petrie, M., Doums, C. & Møller, A. P. 1998. The degree of extra-pair paternity increases with genetic variability. *Proc.Natl.Acad.Sci., USA* **95**: 9390-9395.

Petrie, M., Halliday, T. & Sanders, C. 1991. Peahens prefer peacocks with elaborate trains. *Anim.Behav.* **41**: 323-331.

Potts, W. K., Manning, C. J. & Wakeland, E. K. 1991. Mating patterns in seminatural populations of mice influenced by MHC genotype. *Nature* **352**: 619-621.

Potts, W. K. & Wakeland, E. K. 1993. Evolution of MHC genetic diversity: a tale of incest, pestilence and sexual preference. *Trends in Genetics* **9**: 408-412.

Presciuttini, S., Toni, C., Tempestini, E., Verdiani, S., Casarino, L., Spinetti, I., Stefano, F. D., Domenici, R. & Bailey-Wilson, J. E. 2002. Inferring relationships between pairs of individuals from locus heterozygosities. *BMC Genetics* **3**: 23.

Pusey, A. & Wolf, M. 1996. Inbreeding avoidance in animals. *Trend.Ecol.Evol.* **11**: 201-206.

Puurtinen, M., Ketola, T. & Kotiaho, J. S. 2005. Genetic compatibility and sexual selection. *Trend.Ecol.Evol.* **20**: 157-8.

Queller, D. C. & Goodnight, K. F. 1989. Estimating relatedness using genetic markers. *Evolution* **43**: 258-275.

Roberts, S. C. & Gosling, L. M. 2003. Genetic similarity and quality interact in mate choice decisions by female mice. *Nat.Genet.* **35**: 103-106.

Roberts, S. C., Little, A. C., Gosling, L. M., Perrett, D. I., Carter, V., Jones, B. C., Penton-Voak, I. & Petrie, M. 2005. MHC-heterozygosity and human facial attractiveness. *Evol.Hum.Behav.* **26**: 213-226.

Roberts, S. C., Little, A. C., Gosling, L. M., Jones, B. C., Perrett, D. I., Carter, V., & Petrie, M. 2005. MHC-assortative facial preferences in humans. *Biol.Letters* (*in press*).

Sauermann, U., Nürnberg, P., Bercovitch, F. B., Berard, J. D., Trefilov, A., Widdig, A., Kessler, M., Schmidtke, J. & Krawczak, M. 2001. Increased reproductive success of MHC class II heterozygous males among free-ranging rhesus macaques. *Hum.Genet.* **108**: 249-254.

Seddon, N., Amos, W., Mulder, R. A. & Tobias, J. A. 2004. Male heterozygosity predicts territory size, song structure and reproductive success in a cooperatively breeding bird. *Proc.R.Soc.Lond. B* **271**: 1823-1829.

Sokal, R. R. & Rohlf, F. J. 1995. *Biometry*. New York: Freeman.

Thornhill, R., Gangestad, S. W., Miller, R., Scheyd, G., McCullough, J. K. & Franklin, M. 2003. Major histocompatibility genes, symmetry and body scent attractiveness in men and women. *Behav.Ecol.* **14**: 668-678.

Tregenza, T. & Wedell, N. 2000. Genetic compatibility, mate choice and patterns of parentage: Invited review. *Molecular Ecology* **9**: 1013-1027.

Trivers, R. L. 1972. Parental investment and sexual selection. In *Sexual Selection and the Descent of Man, 1871-1971* (ed. B. Campbell), pp. 136-179. London: Heinemann Press.

Wedekind, C. & Furi, S. 1997. Body odour preferences in men and women: do they aim for specific MHC combinations or simply heterozygosity? *Proc.R.Soc.Lond. B* **264**: 1471-79.

Wedekind, C., Seebeck, T., Bettens, F. & Paepke, A. J. 1995. MHC-dependent mate preferences in humans. *Proc.R.Soc.Lond. B* **260**: 245-249.

Yamazaki, K., Boyse, E. A., Mike, V., Thaler, H. T., Mathieson, B. J., Abbott, J., Boyse, J., Zayas, Z. A. & Thomas, L. 1976. Control of mating preferences in mice by genes in the major histocompatability complex. *J.Exp.Med.* **144**: 1324-1335.

Zeh, J. A. & Zeh, D. W. 1996. The evolution of polyandry I: Intragenomic conflict and genetic incompatibility. *Proc.R.Soc.Lond. B* **263**: 1711-1717.

Zeh, J. A. & Zeh, D. W. 2003. Toward a new sexual selection paradigm: Polyandry, conflict and incompatibility. *Ethology* **109**: 929-950.

Figure 1. Theoretical relationship between heterozygosity and number of shared alleles across four diallelic loci in Hardy-Weinberg equilibrium. Allele frequencies at all four loci were the same within each polymorphism scenario. The higher the level of polymorphism (heterozygosity expected under Hardy-Weinberg equilibrium; H_E), the more positive the relationship is. Allele frequencies for the five scenarios are: $p = 0.9$, $q = 0.1$ ($H_E = 0.18$), $p = 0.8$, $q = 0.2$ ($H_E = 0.32$), $p = 0.7$, $q = 0.3$ ($H_E = 0.42$), $p = 0.6$, $q = 0.4$ ($H_E = 0.48$), and $p = 0.5$, $q = 0.5$ ($H_E = 0.50$). At very unequal allele frequencies (i.e. $p = 0.9$, $q = 0.1$) the relationship between heterozygosity and shared alleles is negative, while at very even allele frequencies (e.g. $p = 0.5$, $q = 0.5$) the relationship is positive.
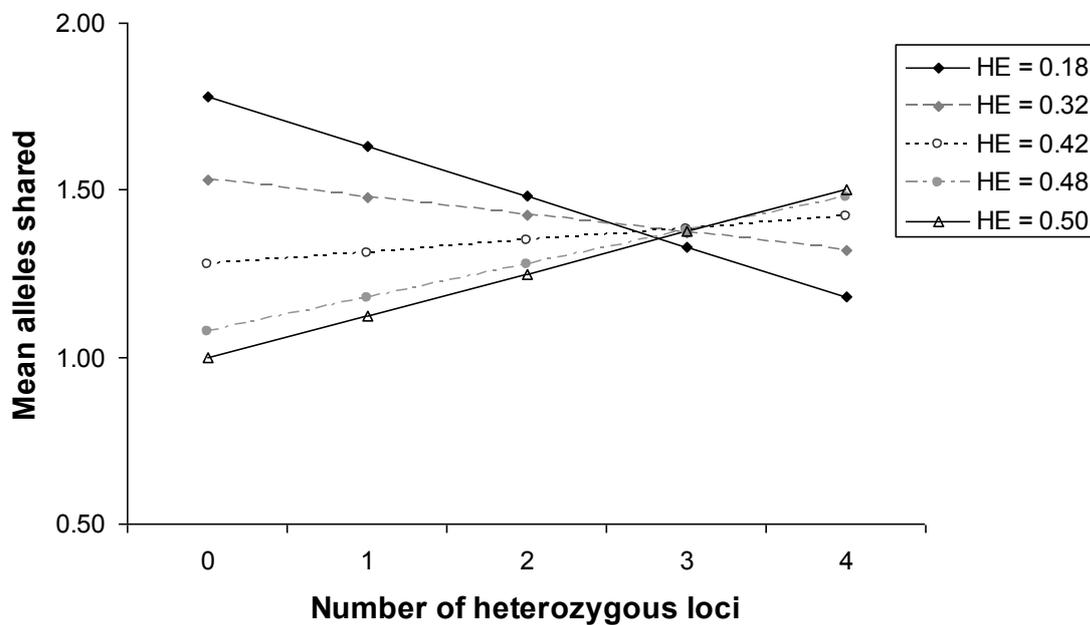
Figure 2. Observed relationships between individual heterozygosity at HLA loci and measures of average genetic similarity in humans. (a) Individuals who are heterozygous at all measured loci share more alleles, on average, with other members of the population. (b) Mean relatedness is negatively related with individual heterozygosity. Data show mean (± standard error) number of shared alleles or relatedness scores for 441 men and women who are heterozygous at 0-3 of 3 HLA loci (N=1, 21, 105 and 314, respectively).
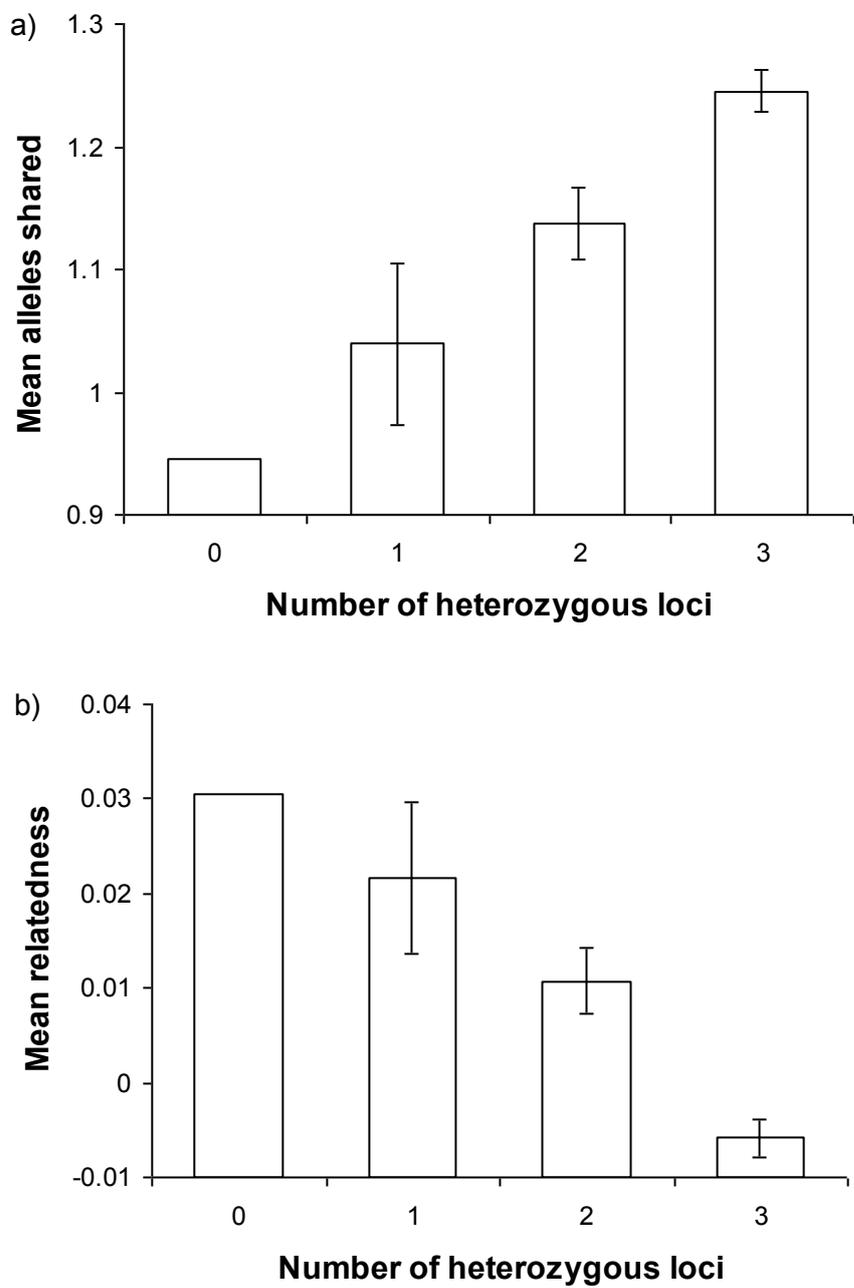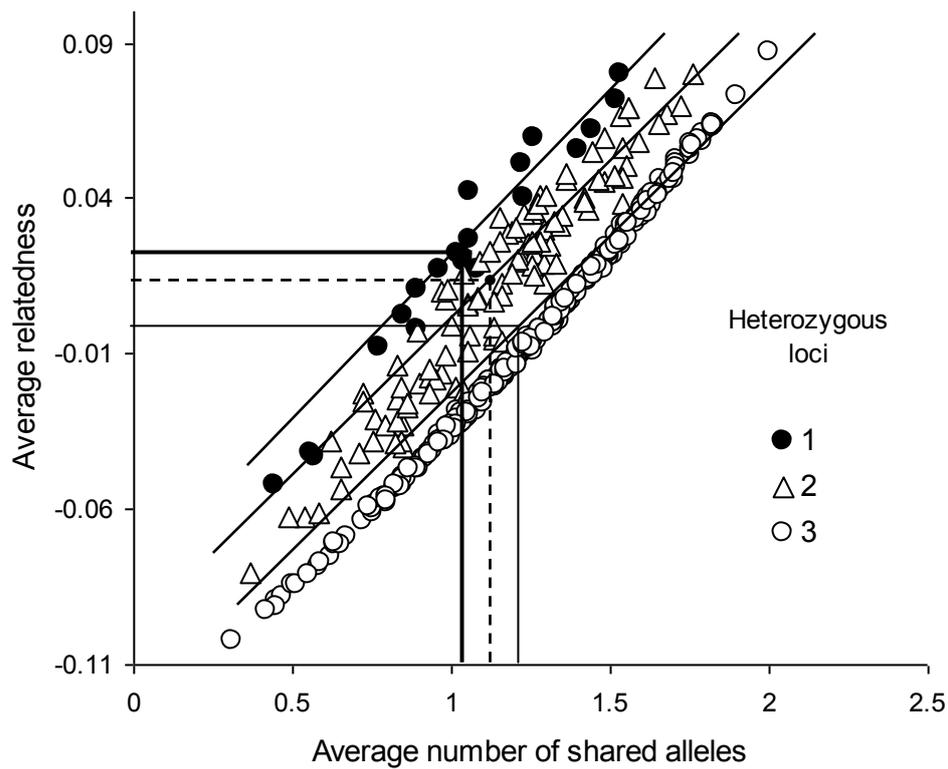
Figure 3. Correlations between average relatedness and average allele sharing at HLA loci in human individuals of variable heterozygosity. Lines of best-fit are shown for each degree of heterozygosity. For each measure of similarity, the bold, dotted and solid lines indicate mean similarity scores for the individuals heterozygous at one, two or three loci, respectively (see also Fig.2).

Relationship between heterozygosity and genetic similarity in peafowl calculated as (*a*) the number of shared alleles and (*b*) Queller & Goodnight's relatedness, for groups of microsatellite loci with different polymorphism levels, for actual matings that took place within the population. 'Low poly' refers to a subset of 4 loci with an average polymorphism (H_E) of 0.15. 'Mod poly' refers to a subset of moderately polymorphic loci with an average polymorphism (H_E) of 0.56. No significant relationships between heterozygosity and genetic similarity existed for the moderately polymorphic set of loci, yet for the low polymorphism loci there was a very strong negative relationship between heterozygosity and genetic similarity measured as both the number of shared alleles and relatedness. Error bars are the standard error of the mean, *n* = 388.
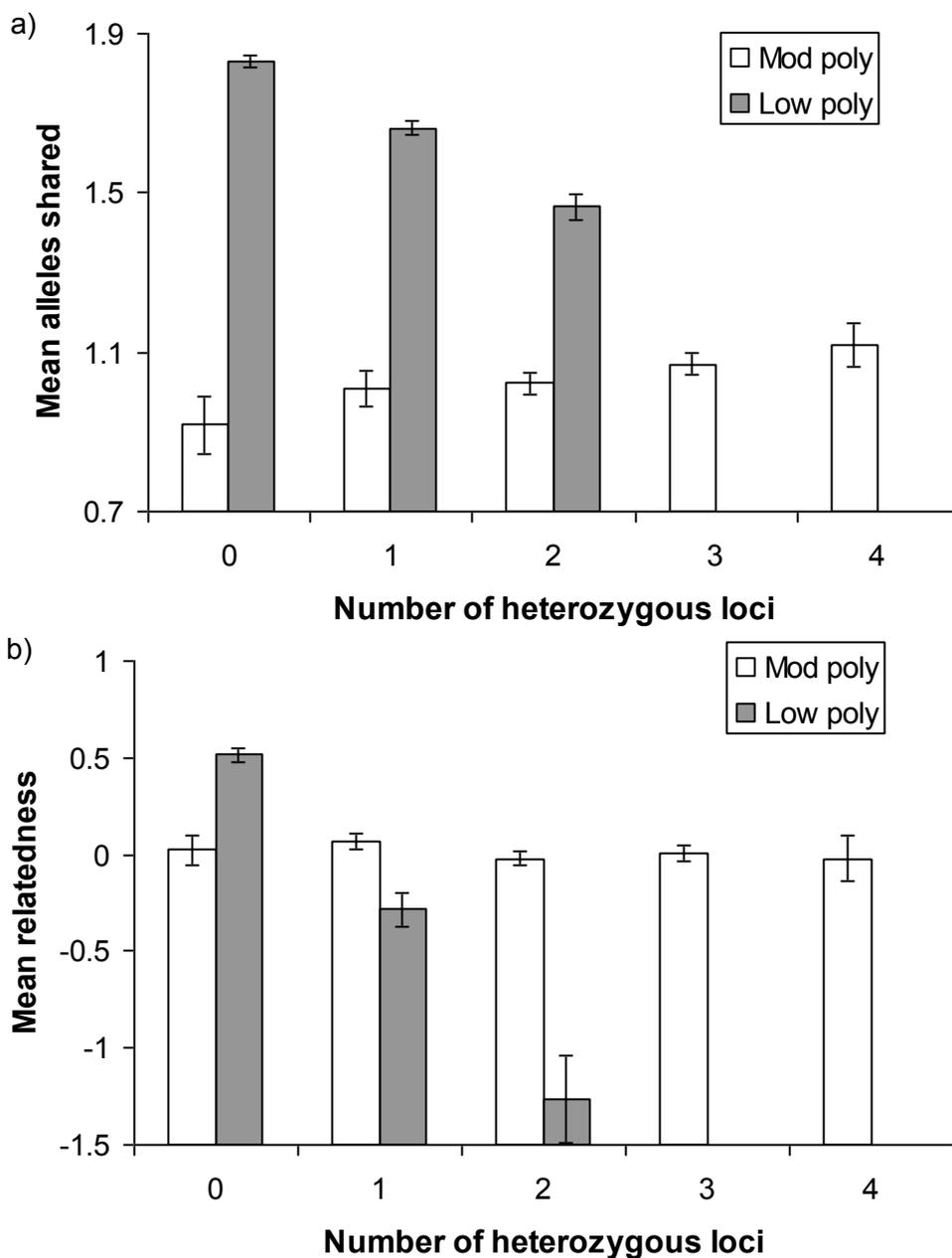
Figure 5. The relationship between relatedness and heterozygosity across a range of polymorphisms, showing that increasing heterozygosity always results in lower relatedness where shared alleles are held constant. 1 Het, 2 Het and 3 Het refer to individuals with 1, 2 or 3 heterozygous loci.