

Correlation between the individual heterozygosity of parents and their offspring

J. B. MITTON*, W. S. F. SCHUSTER†, E. G. COTHRAN‡ & J. C. DE FRIES§

*Department of Environmental, Population and Organismic Biology, Campus Box 334, University of Colorado, Boulder, CO 80309, †Black Rock Forest, Box 483, Cornwall, NY 12518, ‡Veterinary Science, University of Kentucky, Lexington, KY 40546-0076 and §Institute for Behavioral Genetics, University of Colorado, Boulder, CO 80309, U.S.A.

Specific formulations are derived for the correlation between the heterozygosity of a randomly mating parent and its offspring for a diallelic locus, and for the correlation when multiple loci are considered. The expected correlation is maximal, approaching $r = 0.50$, when allelic frequencies are highly asymmetric, and it is zero when the allelic frequencies are equal to 0.50. Parent–offspring correlations, up to a maximum of 0.50 for diallelic loci, indicate that levels of heterozygosity can respond to selection. Multilocus allozyme data from limber pine, *Pinus flexilis*, and from horses of standardbred and thoroughbred breeds are used to demonstrate correlations between a parent and its offspring. The Spearman rank correlation between the heterozygosity of a limber pine and the mean heterozygosity of her offspring is $r = 0.45$. Correlations in the horses range from $r = 0.16$ to 0.32.

Keywords: heterozygosity, horses, limber pine, parent–offspring correlations.

Introduction

Two generalizations characterize the pattern of transmission of genotypes at diallelic loci within randomly mating populations. Regardless of the allelic frequencies, half of the heterozygous offspring have homozygous mothers, and half of them have heterozygous mothers. Regardless of the genotype(s) of the father or fathers, heterozygous mothers produce 50 per cent heterozygous offspring. Here we add a third generalization by describing the expected correlation between the heterozygosity of randomly mating parents and their offspring. We derive a specific formulation for this correlation as a function of allelic frequencies, and we report correlations of individual heterozygosity between parents and their offspring in limber pine, *Pinus flexilis*, and from pedigreed horses of standardbred and thoroughbred breeds.

Theoretical expectations

A correlation between individual heterozygosities of parents and their offspring arises from the fact that, at most allelic frequencies, heterozygous parents produce

higher proportions of heterozygous progeny than do homozygous parents. Consider a locus segregating 2 alleles, A and a , with frequencies p and q , respectively, in an infinite population with random mating and no selection. To simplify the exposition, we will consider the production of progeny genotypes from the perspective of the maternal parent. Homozygous mothers produce 50 per cent heterozygous offspring when allelic frequencies are $p = q = 0.5$, but at all other frequencies they produce lower proportions of heterozygous offspring. More precisely, the proportion of heterozygous progeny from heterozygous mothers is $pq/(2pq)$ while the proportion of heterozygous progeny from homozygous mothers is $pq/(p^2 + q^2)$. As allelic frequencies become increasingly unequal, the proportion of heterozygous progeny produced by homozygous mothers decreases.

The expected value for the correlation between offspring and parent heterozygosity, r_{op} , can be derived most simply from the formulations of genotypic value, additive genetic variance, and dominance genetic variance (Falconer, 1989). Homozygous genotypes are assigned genotypic values of a and $-a$, and the heterozygote is assigned a value d . When homozygous and heterozygous genotypes are assigned the scores of 0 and 1, respectively, the genotypic values of a and d

*Correspondence.

become 0 and 1, respectively. Additive genetic variance is

$$V_A = 2pq\alpha^2 \text{ where } \alpha = a + d(q-p) = q-p$$

and therefore

$$V_A = 2pq(q-p)^2.$$

Dominance variance is

$$V_D = (2pqd)^2 = 4p^2q^2.$$

The correlation of offspring heterozygosity with parental heterozygosity, r_{op} , is

$$r_{op} = 1/2V_A/(V_A + V_D) = pq(q-p)^2/[2pq(q-p)^2 + 4p^2q^2] \\ = (q-p)^2/[2(q-p)^2 + 4pq].$$

Figure 1 presents r_{op} as a function of q , and shows that the correlation is equal to zero only when $p = q = 0.5$, and that it increases to a maximum of $r_{op} = 0.5$ as allelic frequencies become increasingly unequal.

The expected correlation can be extended to M independent loci by recalling that V_A and V_D can each be summed across loci as follows:

$$V_A = V_{A1} + V_{A2} + \dots + V_{AM}$$

and

$$V_D = V_{D1} + V_{D2} + \dots + V_{DM}.$$

Therefore,

$$r_{op} = \frac{\sum_{i=1}^M p_i q_i (q_i - p_i)^2}{2 \sum_{i=1}^M p_i q_i (q_i - p_i)^2 + 4 \sum_{i=1}^M p_i^2 q_i^2}$$

where p_i is assigned to the allele with the higher frequency. Examination of a series of numerical examples revealed that a reasonable approximation of this correlation can be achieved by using the mean allelic frequency and the single locus expression for r_{op} ; this approximation was usually within 0.02 of the multilocus r_{op} . The deviations are probably due to the non-linear relationship between p and r_{op} .

To demonstrate the generality of correlations between parents and their offspring, multilocus correlations are demonstrated in a natural population of limber pine and in standardbred and thoroughbred horses. The observed correlations are similar in sign and magnitude with the expected correlations.

Materials and methods

Limber pine allozyme data are taken from a paternity analysis of an isolated stand of limber pine on the Pawnee National Grasslands, in northeastern Colorado. Individual heterozygosities were based on

the following 10 polymorphic allozyme loci: phosphoglucosyltransferase-2, uridine-5-diphosphoglucose pyrophosphorylase-1 and 2, 6-phosphoglucosyltransferase-1 and 2, malate dehydrogenase, phosphoglucose isomerase, shikimate dehydrogenase, peroxidase, and fluorescent esterase (Schuster, 1989). All the loci in limber pine fit Hardy-Weinberg expectations. The Spearman rank correlation coefficient was used to test the null hypothesis of no association between parent and offspring heterozygosities. A total of 467 seeds was collected from the cones of 69 maternal trees, and data were analysed both as observations of maternal heterozygosity paired with the heterozygosity of one paired seed, and also with the mean heterozygosity of all seeds.

Allozyme data of horses are from samples submitted to the University of Kentucky Equine Blood Typing Research Laboratory for genetic typing for parentage verification prior to registration. Only parent-offspring combinations where neither sire nor dam could be excluded as the parents of the foal were used. The loci analysed for all individuals were transferrin, a 1- β glycoprotein, albumin, serum esterase, and 6-phosphoglucosyltransferase. The remaining loci are vitamin D-binding protein, α -haemoglobin, glucose-phosphate isomerase, phosphoglucosyltransferase, and the protease inhibitor system. All the genetic data in horses except transferrin in standardbreds fit Hardy-Weinberg expectations. Genetic analyses are based on 6-10 allozyme loci, with a mean of 9.6. The numbers vary because some loci were not examined when the samples were processed. Only individuals with data for at least six of these loci were included in the analysis. To accommodate the variable number of loci, the data were transformed to the proportion of loci heterozygous for an individual. Any data point, composed of the proportion of loci heterozygous for any two individuals, was trimmed first to the set of loci scored in both individuals. The Spearman rank correlation coefficient was used to test the null hypothesis of no correlation of heterozygosities in comparisons of foals and sires, foals and dams, sires and dams, and among randomly chosen foals.

For each of the data sets, a simulation program was run to determine the expected correlation between parental and offspring heterozygosity. These simulations used the number of loci, the numbers of alleles, and the allelic frequencies in that particular data set. All loci were assumed to be independently assorting, and mating was random with respect to the genes of interest. For the data from limber pine, the simulation was run two ways, first to estimate the correlation between a maternal parent and a single seedling, then to estimate the correlation between a maternal parent

and the mean of her seedlings. A population of 100 individuals was established with the aid of a random number generator by sampling a gene pool with the number of loci, the number of alleles, and the allelic frequencies observed in the population of limber pine. Pollen haplotypes were established the same way. Then, for each individual in the population, a randomly chosen gamete was extracted and joined with a randomly chosen pollen haplotype, producing one offspring from each individual in the population. The correlation between maternal and offspring heterozygosity was then calculated across the 100 parent-offspring pairs. This simulation was run 100 times to estimate the mean and the standard deviation of the correlation. The simulation to estimate the correlation between a maternal parent and the mean of her offspring established a pool of maternal parents as described above. However, rather than a single offspring for each parent, 10 offspring were produced by each parent; the correlation was then calculated between maternal heterozygosity and the mean heterozygosity of 10 offspring. In this simulation, each pollen haplotype was sampled independently, so that the offspring represent half sibs. This mating system is most appropriate for the population of limber pine in this study (Schuster, 1989).

Results

The individual heterozygosity of embryos in limber pine seed is correlated with the heterozygosity of the trees that bear them. When the correlation is computed with maternal heterozygosity and the heterozygosity of one of her embryos, the Spearman rank correlation coefficient is $r_s = 0.22$ ($N = 467$, $P < 0.001$). When the relationship is tested between maternal heterozygosity and the mean heterozygosity of her embryos, the correlation is $r_s = 0.45$ ($N = 69$, $P < 0.001$). In the simulations, the mean correlation for single offspring was $r = 0.15$ (s.d. = 0.104), and the mean correlation for families was $r = 0.36$ (s.d. = 0.073).

In standardbred and thoroughbred horses, the heterozygosity of foals is positively correlated with the heterozygosity of both the sires and dams (Table 1). The correlation coefficients range from a low of $r_s = 0.16$ ($P < 0.001$) between foals and sires in standardbreds to $r_s = 0.32$ ($P < 0.001$) between foals and their sires in thoroughbreds. Heterozygosities of sires and dams were uncorrelated, and the heterozygosities of randomly chosen foals were uncorrelated. In simulations run with the allelic frequencies observed in standardbred horses, the mean of the correlations was $r = 0.13$ (s.d. = 0.04). The correlations expected between sires and their foals ranged from 0.11 to 0.12,

Table 1 Correlations between individual heterozygosities in standardbred and thoroughbred horses

| | Standardbred | | | Thoroughbred | | | r_{exp} |
|--------------|--------------|------|-----|--------------|-----|-----|------------------|
| | r | N | P | r | N | P | |
| Foal, sire | 0.16 | 3351 | *** | 0.32 | 263 | *** | 0.12 |
| Foal, dam | 0.20 | 3474 | *** | 0.22 | 263 | *** | 0.16 |
| Sire, dam | -0.01 | 2599 | NS | 0.06 | 263 | NS | |
| Random foals | 0.02 | 4255 | NS | 0.05 | 263 | NS | |

Note: r is the correlation coefficient; N is sample size; ***Probability < 0.001 ; NS is non-significant.

r_{exp} is the mean of the correlation from computer simulations. See text for further explanation.

while the correlations expected between dams and their foals ranged from 0.16 to 0.20.

Discussion

Although the correlations between heterozygosity of parents and offspring have been illustrated here with allozyme loci, they will be found at all diploid, nuclear loci; the relationship in Fig. 1 is general. The correlation between the heterozygosities of parents and their offspring indicates that heterozygosity can respond to selection, to a maximum frequency of 0.50 for diallelic loci.

A correlation between maternal and offspring heterozygosity has been reported in white-tailed deer (*Odocoileus virginianus*) (Cothran *et al.*, 1983). Individual heterozygosity was estimated with four polymorphic protein loci, and in a sample of 153 mothers and their offspring, offspring heterozygosity increased with maternal heterozygosity ($r = 0.37$, $P < 0.001$). Because both twinning and embryonic growth rate increase with allozyme heterozygosity in white-tailed deer, and because offspring and maternal heterozygosities are correlated, twinning and foetal growth rate may vary among families.

Theoretical studies of fitness determination suggest that, at loci whose variation is maintained by balancing selection, fitness increases with heterozygosity (Ginzburg, 1979, 1983; Turelli & Ginzburg, 1983). This general expectation has provided the incentive for numerous empirical studies of the relationship between individual heterozygosity, the number of loci from a sample for which an individual is heterozygous (Mitton & Pierce, 1980), and various components of fitness (Zouros *et al.*, 1980; Koehn & Shumway, 1982; Pierce & Mitton, 1982; Garton, 1984; Hawkins *et al.*, 1986, 1989; Rodhouse *et al.*, 1986; Bush *et al.*, 1987;

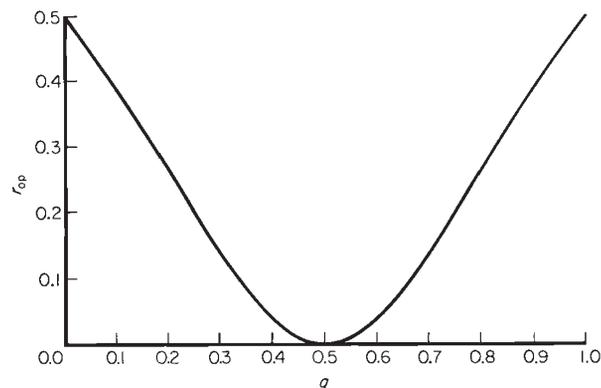


Fig. 1 Correlation between individual heterozygosities of offspring and their parents r_{op} as a function of allelic frequency, q . When p and q are allelic frequencies at a diallelic locus, $r_{op} = (q - p)^2 / [2(q - p)^2 + 4pq]$. See text for further explanation.

Danzmann *et al.*, 1987; Diehl, 1988; Koehn *et al.*, 1988; Gajardo & Beardmore, 1989; Ferguson & Drahushchak, 1990; Teska *et al.*, 1990; Mopper *et al.*, 1991; Pecon Slattery *et al.*, 1991). These studies reveal that fitness generally increases with heterozygosity (see reviews by Beardmore, 1983; Mitton & Grant, 1984; Allendorf & Leary, 1986; Ledig, 1986; Zouros & Foltz, 1987; Mitton, 1989, 1993). Of course, if overdominance were present at a locus that influences relative fitness, then r_{op} for relative fitness would equal zero at equilibrium, when

$$q = s_1 / [s_1 + s_2],$$

where s_1 and s_2 are the coefficients of selection against the two homozygotes. Thus, r_{op} for individual heterozygosity, which is zero only when $p = q = 0.5$, should exceed that for relative fitness at or near selection equilibrium. However, it is very unlikely that endangered populations, population facing increasing habitat destruction and pressure from humans, and populations stressed by climatic change, are at equilibrium. In these situations, fitness differentials associated with heterozygosity could be heritable.

Acknowledgements

This work was begun while J. B. M. was on fellowship at Stanford University's Hopkins Marine Station. We acknowledge insight and helpful suggestions during conversations with Dennis Powers, Marc Feldman and Freddy Christiansen. Michael Grant, Pat Carter and Anna Goebel contributed comments on the manuscript, and Hsiu-Ping Liu prepared the figure. Partial support was provided from a faculty fellowship awarded to J. B. M. by the Committee on Research and Creative Work, University of Colorado.

References

- ALLENDORF, F. W. AND LEARY, R. F. 1986. Heterozygosity and fitness in natural populations of animals. In: Soule, M. E. (ed.) *Conservation Biology: The Science of Scarcity and Diversity*, Sinauer Associates, Sunderland, MA, pp. 57-76.
- BEARDMORE, J. A. 1983. Extinction, survival, and genetic variation. In: Schonewald-Cox, C. M., Chambers, S. M., MacBryde, B. and Thomas, W. L. (eds), *Genetics and Conservation: A Reference for Managing Wild Animal and Plant Populations*, Benjamin/Cummings Publishing Co., Menlo Park, CA, pp. 125-151.
- BUSH, R. M., SMOUSE, P. E. AND LEDIG, F. T. 1987. The fitness consequences of multiple-locus heterozygosity: the relationship between heterozygosity and growth rate in pitch pine (*Pinus rigida* Mill.). *Evolution*, **41**, 787-798.
- COTHRAN, E. G., CHESSER, R., SMITH, M. H. AND JOHNS, P. E. 1983. Influences of genetic variability and maternal factors on fetal growth in white-tailed deer. *Evolution*, **37**, 282-291.
- DANZMANN, R. G., FERGUSON, M. M. AND ALLENDORF, F. W. 1987. Heterozygosity and oxygen-consumption rates as predictors of growth and developmental rate in rainbow trout. *Physiol. Zool.*, **60**, 211-220.
- DIEHL, W. J. 1988. Genetics of carbohydrate metabolism and growth in *Eisenia foetida* (Oligochaeta: Lumbricidae). *Heredity*, **61**, 379-387.
- FALCONER, D. S. 1989. *Introduction to Quantitative Genetics*. 3rd edn, John Wiley and Sons, New York.
- FERGUSON, M. M. AND DRAHUSHCHAK, L. R. 1990. Disease resistance and enzyme heterozygosity in rainbow trout. *Heredity*, **64**, 413-417.
- GAJARDO, G. M. AND BEARDMORE, J. A. 1989. Ability to switch reproductive mode in *Artemia* is related to maternal heterozygosity. *Marine Ecol. Prog. Ser.*, **55**, 191-195.
- GARTON, D. W. 1984. Relationship between multiple locus heterozygosity and physiological energetics of growth in the estuarine gastropod *Thais haemastoma*. *Physiol. Zool.*, **57**, 530-543.
- GINZBURG, L. R. 1979. Why are heterozygotes often superior in fitness? *Theoret. Popul. Biol.*, **15**, 264-267.
- GINZBURG, L. R. 1983. *Theory of Natural Selection and Population Growth*. Benjamin/Cummings, Menlo Park, CA.
- HAWKINS, A. J. S., BAYNE, B. L. AND DAY, A. J. 1986. Protein turnover, physiological energetics and heterozygosity in the blue mussel, *Mytilus edulis*: the basis of variable age-specific growth. *Proc. R. Soc. Lond. B*, **229**, 161-176.
- HAWKINS, A. J. S., BAYNE, B. L., DAY, A. J., RUSIN, J. AND WORRALL, C. M. 1989. Genotype-dependent interrelations between energy metabolism, protein metabolism and fitness. In: Ryland, J. S. and Tyler, P. A. (eds), *Reproduction, Genetics and Distributions of Marine Organisms*. Olsen and Olsen, Fredensborg, Denmark, pp. 283-292.
- KOEHN, R. K., DIEHL, W. J. AND SCOTT, T. M. 1988. The differential contribution by individual enzymes of glycolysis and protein catabolism to the relationship between heterozygosity and growth rate in the coot clam, *Mulinia lateralis*. *Genetics*, **118**, 121-130.
- KOEHN, R. K. AND SHUMWAY, S. E. 1982. A genetic/physiological explanation for differential growth rate among individuals

- of the American oyster, *Crassostrea virginica* (Gmelin). *Marine Biol. Lett.*, **3**, 35-42.
- LEDIG, F. T. 1986. Heterozygosity, heterosis, and fitness in outbreeding plants. In: Soule, M. E. (ed.), *Conservation Biology: The Science of Scarcity and Diversity*, Sinauer Associates, Sunderland, MA, pp. 77-104.
- MITTON, J. B. 1989. Physiological and demographic variation associated with allozyme variation. In: Soltis, D. and Soltis, P. (eds) *Isozymes in Plants*, Dioscorides Press, Portland, OR, pp. 127-145.
- MITTON, J. B. 1993. Theory and data pertinent to the relationship between heterozygosity and fitness. In: Thornhill, N. (ed.) *The Natural History of Inbreeding and Outbreeding*, University of Chicago Press, Chicago, IL, pp. 17-41.
- MITTON, J. B. AND GRANT, M. C. 1984. Associations among protein heterozygosity, growth rate, and developmental homeostasis. *Ann. Rev. Ecol. Syst.*, **15**, 479-499.
- MITTON, J. B. AND JEFFERS, R. M. 1989. The genetic consequences of mass selection for growth rate in Engelmann spruce. *Silvae Genet.*, **38**, 6-12.
- MITTON, J. B. AND PIERCE, B. A. 1980. The distribution of individual heterozygosity in natural populations. *Genetics*, **95**, 1043-1054.
- MOPPER, S., MITTON, J. B., WHITHAM, T. G., COBB, N. S. AND CHRISTENSEN, K. M. 1991. Genetic differentiation and heterozygosity in pinyon pine associated with resistance to herbivory and environmental stress. *Evolution*, **45**, 989-999.
- PECON SLATTERY, J., VRIJENHOEK, R. C. AND LUTZ, R. A. 1991. Heterozygosity, growth, and survival of the hard clam, *Mercenaria mercenaria*, in seagrass vs sandflat habitats. *Marine Biol.*, **111**, 335-342.
- PIERCE, B. A. AND MITTON, J. B. 1982. Allozyme heterozygosity and growth in the tiger salamander, *Ambystoma tigrinum*. *J. Hered.*, **73**, 250-253.
- RODHOUSE, P. G., MCDONALD, J. H., NEWELL, R. I. E. AND KOEHN, R. K. 1986. Gamete production, somatic growth and multiple locus heterozygosity in *Mytilus edulis* L. *Marine Biol.*, **90**, 209-214.
- SCHUSTER, W. S. F. 1989. Population structure and demography of an isolated stand of limber pine (*Pinus flexilis* James). Ph. D Thesis, University of Colorado, Boulder, CO.
- TESKA, W. R., SMITH, M. H. AND NOVAK, J. M. 1990. Food quality, heterozygosity, and fitness correlated in *Peromyscus polionotus*. *Evolution*, **44**, 1318-1325.
- TURELLI, M. AND GINZBURG, L. 1983. Should individual fitness increase with heterozygosity? *Genetics*, **104**, 191-209.
- ZOUROS, E. AND FOLTZ, D. W. 1987. The use of allelic isozyme variation for the study of heterosis. In: Rattazzi, M. C., Scandalios, J. G. and Whitt, G. S. (eds), *Isozymes: Current Topics in Biological and Medical Research*, Vol. 13, Alan R. Liss, New York, pp. 1-59.
- ZOUROS, E., SINGH, S. M. AND MILES, H. E. 1980. Growth rate in oysters: an overdominant phenotype and possible explanations. *Evolution*, **34**, 856-867.