Zoo Biology 5:171–179 (1986)

Considerations on the Conservation of Alleles and of Genic Heterozygosity in Small Managed Populations

Paul A. Fuerst and Takeo Maruyama

Department of Genetics, The Ohio State University, Columbus, Ohio (P.A.F.); National Institute of Genetics, Mishima, Japan (T.M.)

Data has been presented on the loss of alleles from populations of small fixed size. Emphasis has been placed on the effect which the allele frequency distribution and the interlocus distribution of heterozygosity will have on the pattern of allelic loss. Rare alleles are rapidly lost during the initial sampling period, and continue to be lost for several generations following the establishment of small population size. The nonequilibrium nature of the process of loss of genetic variability is stressed. The rapid loss of rare alleles and the preservation of intermediate and high frequency alleles will result in (1) heterozygosity declining much more slowly than allele number, and (2) the establishment of genetically similar populations when sampled from the same base population.

Key words: allelic diversity, nonequilibrium population, population bottleneck

INTRODUCTION

Genetic considerations about the management of endangered species have emphasized the preservation of genetic variability. Difficulties in maintaining a viable population can occur when the population tends towards genetic uniformity; inbreeding in a previously outbred population produces deleterious homozygotes concurrently with the loss of heterozygotes. Both effects can be serious. Homozygosity for a disadvantageous allele is obviously detrimental, while it has often been suggested that heterozygosity, per se, can be advantageous, and that the effect is independent of any relationship with rare recessive deleterious alleles [Beardmore, 1983]. The rate of loss of genic heterozygosity from a population with a defined effective size is well known, and has been presented in many contexts since it was first determined correctly by Sewall Wright. The rate of loss of alleles from the population is less well known, although several approaches have been used in attempts to estimate the allelic loss following a sharp reduction in the population size. Such loss will occur whenever a sample of an endangered population is taken into a managed environment.

Received September 17, 1985; accepted October 22, 1985.

Address reprint requests to Dr. Paul A. Fuerst, Department of Genetcis, Ohio State University, 484 W. 12th Avenue, Columbus, Ohio 43210.

We have recently been studying the changes in both the allelic and the genic variability which follow major changes in population sizes. We have been especially interested in those rapid reductions of the population to a small size which have been termed population bottlenecks [Maruyama and Fuerst, 1984, 1985]. Our studies yield formulas which can be used to predict the patterns of change in the number of alleles in the population. In this note, we will summarize some of our findings, present results specific to the short-term management of animal populations, and will further show that changes in genetic variability should be interpreted in light of the frequency distributions of alleles and of heterozygosity over different loci. These distributions have generally been ignored during considerations of conservation of genetic variability, but they provide some useful insights about the problems facing the population management team.

FACTORS CONTRIBUTING TO LOSS OF ALLELES FROM THE POPULATION

The loss of allelic variability during species management occurs in several ways. First, alleles which are present in a large (or possibly small) natural population undergo an initial sampling during the process of selecting those animals which will make up the managed population. This sampling event will occur whether the managed individuals form a "wild" population within a game reserve or park, or are a group of animals assembled for captive breeding purposes in a zoo. For many species, such as the large primates, this sampling process must be viewed as essentially completed at the present time. This is because of international agreements which restrict the importation of additional specimens from the wild. For other species, the sampling process may not have even begun.

How much of the pre-existing variation is lost to the future managed genepool will differ from species to species. The population geneticist is capable of providing only the broadest, most intuitive guidelines to the species manager on this point. We can obviously advise that large samples from many different sources are better than either small samples or samples from single sources. The management of plant germplasm conservation is considerably ahead of animal gene conservation on such topics, and zoo personnel stand to profit from some of the considerations of the plant conservationist. However, studies of sample sizes necessary for highly effective conservation of rare alleles in the species as a whole suggest that the sample sizes necessary to ensure the preservation of specific alleles will be well beyond the capacity of animal conservation resources [Chapman, 1984]. A large proportion of the natural allelic variability may be lost during this initial process, depending upon the size of the group taken from the wild and degree of geographical sampling from differentiated populations.

Following the initial collection stage, alleles are further lost due to the small finite size of the breeding population. This is the classical process called genetic drift and is of great importance to the species manager. This is the stage where the manager has a chance to intervene to slow down the loss of genetic variability. It has been shown by Kimura [1955] that the equilibrium rate of loss of alleles from a finite population (where the equilibrium is defined in a population which is stationary in size, and which is losing variability at a constant rate due to genetic drift) will be approximately $n(n - 1)/2N_e$ per generation, where n is the number of alleles remaining in the population at any time point and N_e is the effective size of the

TABLE 1. Number of Alleles at the Average Locus in Various Frequency Ranges for Two Levels of Population Heterozygosity. Theoretical Values Obtained by the Formulation of Kimura and Crow (1964)

	Number of alleles	
Allele frequency	$\bar{H} = 0.05$	$\bar{H} = 0.10$
0.001–0.01	0.119 (8.6%)	0.252 (14.2%)
0.01–0.05	0.087 (6.3%)	0.182 (10.3%)
0.05–0.20	0.081 (5.9%)	0.171 (9.6%)
0.20–0.95	0.217 (15.8%)	0.426 (24.0%)
0.95–1.00	0.875 (63.5%)	0.743 (41.9%)
Total alleles >0.001	1.379	1.774

population. This differs from the loss of gene diversity (heterozygosity), which occurs at a rate of $1/2N_e$ per generation. If the average number of alleles per locus is above 1.6, the rate of loss of alleles will be more rapid than the rate of loss of heterozygosity. When allele numbers fall below this value, the relative rate of allelic loss becomes less than the rate of loss of heterozygosity.

These calculations assume an equilibrium rate of loss of variation. However, such equilibrium conditions would not prevail in the initial generations of captive management, which is exactly the time period during which we are most concerned about allelic loss. To calculate the loss of alleles during this nonequilibrium stage involves some statistical considerations which differ from those involved in studying the loss of heterozygosity. Trivially, for instance, allele number will never approach zero, as does heterozygosity, since all loci possess one allele when monomorphism is reached.

One important factor greatly affecting the rate of loss of alleles in the nonequilibrium (realistic) situation, but often ignored during discussions of the loss of genetic variation in a population is the fact that the allele frequency distribution is not bell-shaped. Often, statistical models are used which make the assumption that all alleles start at the same frequency. This is incorrect in describing the total genetic diversity of a managed population. In reality, the allele frequency distribution has been found to be U- or J-shaped (depending upon the effective population size and the mutation rate). This has a major impact upon the initial rate of loss of alleles from the population due to genetic drift. We are not simply sampling alleles from a multinomial distribution of frequencies, as is sometimes assumed [for instance in some of the cases studied by Denniston, 1978 and Allendorf, 1986]. Because of the real nature of the allele frequency distribution, certain allelic classes are much more likely to be included in a sample taken from the population, while other classes are more likely to be lost from the sample.

The importance of these considerations can be appreciated by examining the frequency distributions in Table 1. This table presents two examples of expected equilibrium allele frequency distributions in populations from which samples might be taken to initiate a managed population. The formulations used to obtain these distributions were originally derived by Kimura and Crow [1964]. The J-shaped nature of the allele frequency distribution is evident in the table. These particular equilibrium distributions represent populations with average heterozygosities of 0.05 and 0.10, values not greatly different from that found in many populations of large mammals [Ryman et al, 1980; Wooten and Smith, 1985]. Chakraborty et al [1980]

have shown that such theoretical distributions are very close to those actually found in natural populations surveyed using protein electrophoresis. If the heterozygosities used to produce the values in Table 1 are not representative of some natural populations of mammals, it is probably because the natural populations have lower genetic variability, not higher. The distributions in these populations would tend to be further J-shaped, with even fewer alleles in intermediate frequencies. Lower average heterozygosity values have, in fact, been suggested by some studies which used two-dimensional electrophoresis to survey a greater number of loci than traditional electrophoretic methods allow. Even if we assume that considerable "hidden" electrophoretic variability is present, because electrophoresis cannot identify all protein differences, there is likely to be little change in the estimated heterozygosity values for large mammals. This is because most hidden variability in populations occurs at the more variable loci which are rarer in mammalian populations than in organisms such as *Drosophila*, and most additional variability involves low frequency alleles, which contribute in only a minor way to heterozygosity [Chakraborty et al, 1980].

The pattern of alleles illustrated in Table 1 is probably close to that found in most natural populations of large animals. Examination of Table 1 shows that there is a large (proportionally) group of alleles which exist at low frequency, and will be easily lost, both during the initiation of the managed population and during the early stages of manipulation, unless specific measures are taken to preserve them. These are principally rare alleles. It is often stated that the rare alleles are likely to be mainly deleterious in nature, and possibly not worth saving as a consequence. The data in Table 1, however, suggest that a substantial fraction of alleles which are normally present in the population will exist at these rare frequency states. The alleles are not deleterious, but rather are selectively similar to those at higher frequency. Loss of these alleles will not, however, be reflected as loss in heterozygosity, because low frequency alleles can be shown to contribute only slightly to the level of heterozygosity. Table 1 also shows that there is a major second group of alleles (high frequency alleles) which will be very hard to lose, even in very small populations. Their high frequency is likely to ensure their inclusion in almost all samples, irrespective of size.

A further consideration related to the distributions in Table 1, which affects our estimates of the loss of allelic variation in a nonequilibrium situation, is the way that the variability is actually distributed over different loci. Unlike the assumptions which are made in many theoretical studies, we must remember that loci in natural populations are not strictly replicates of one another. In fact, there is a well defined distribution of genic heterozygosity at different loci. Information on the theory of this distribution, and an examination of data from a variety of natural populations has been reviewed by Fuerst et al [1977].

The theoretical distribution of heterozygosity for several different values of average heterozygosity is shown in Figure 1. The distributions tend to be L-shaped, suggesting that most loci have little or no variability present, especially when the average heterozygosity of the species is small. Remember, in most species of larger mammals (which make up a large proportion of our managed species) average heterozygosity is low [Ryman et al, 1980; Wooten and Smith, 1985]. The alleles present at those loci which make up the large class of relatively invariant loci will consist of a single high frequency allele, and, possibly, several rare alleles. Such loci will quickly become homozygous as sampling occurs to produce a population for management purposes. This process takes place, again, without regard to the selective

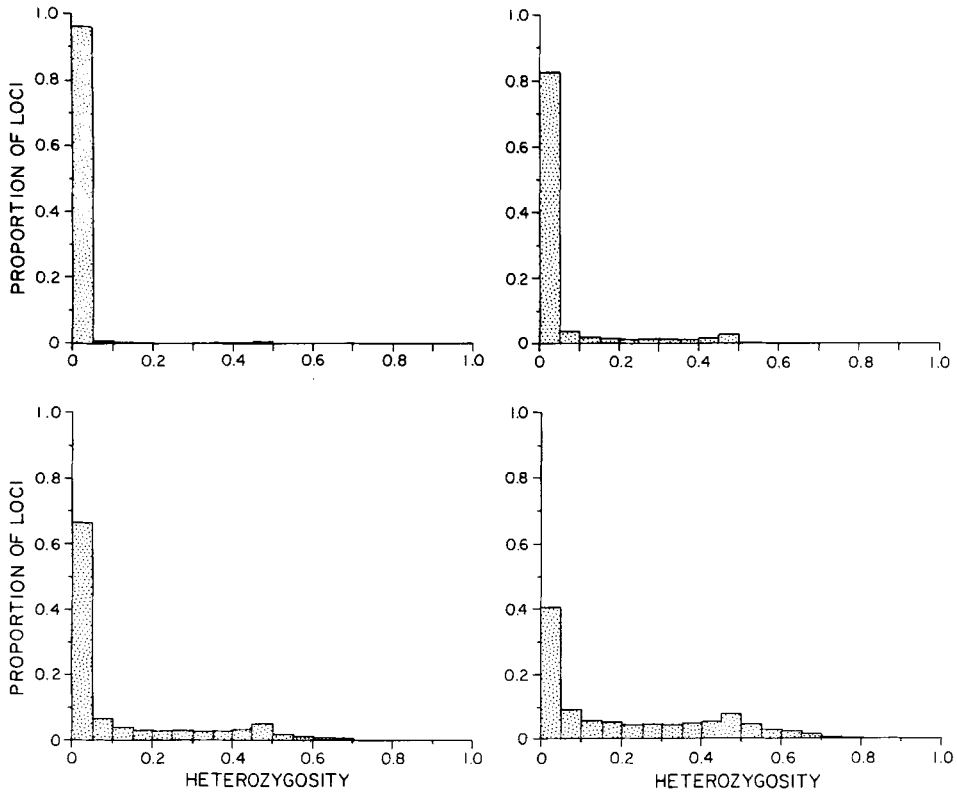

Fig. 1. The distribution of single locus heterozygosities for different values of average heterozygosity. (A, top left) $\overline{H} = 0.10$, (B, top right) $\overline{H} = 0.05$, (C, bottom left) $\overline{H} = 0.10$, (D, bottom right) $\overline{H} = 0.20$.

usefulness of the rare alleles which are being lost. All samples taken from the population are likely to be similar, or even identical, for this entire array of loci, with all variant alleles being lost. These loci, however, will contribute only slightly to any loss of heterozygosity, since rare alleles add little to the increase in average heterozygosity. We are likely, therefore, to lose a significant number of alleles without major changes in heterozygosity.

While most loci represented in Figure 1 have low levels of variation, a minority of loci are highly polymorphic. During the sampling processes, these loci may remain relatively polymorphic, even in the face of a severe population bottleneck, because they contain several intermediate frequency alleles. In fact, we have shown elsewhere that the polymorphic loci are likely to remain polymorphic as the population declines in size or goes through repeated cycles of sampling [Huettle et al, 1980]. Since alleles with intermediate frequencies contribute greatly to heterozygosity, we can have a severe reduction in allele number with only a minor reduction in average heterozygosity. A set of populations which result from this process will be quite homogeneous, with respect to allelic similarity, although they retain significant average heterozygosity. This could have major implications in a situation in which several separate demes had been set up to retain more genetic variability than a single panmictic unit. In fact, the sampling effects mediated by the patterning of variability over loci might result in

an extremely homogeneous population due to loss of low frequency alleles and retention of the same high and intermediate level alleles.

The patterning of heterozygosity over loci has generally been ignored by workers studying the loss of either heterozygosity or alleles from a population. If the decline of genetic variability occurs very gradually under an equilibrium situation this would not be serious. However, it is likely that no managed population is in equilibrium, and the decline in variability is likely to be rapid.

RATE OF LOSS OF ALLELES

Several investigators have studied how allele numbers change under the pressure of sampling. Denniston [1978] studied the loss of alleles theoretically, but without considering the effects caused by the allele frequency distribution. He assumed that alleles started out at equal frequency. This is certainly not the case. Allendorf [1986] examined a similar situation by computer simulation. Both these studies give useful insights about the behavior of particular groups of alleles, but only approximate the overall pattern of allelic decline. The allele frequency distribution is a dynamic measure. Because it depends on the interaction of mutation and genetic drift, it changes as we begin the sampling process, and we must consider a nonequilibrium process during the early stages of establishing a managed population.

We have developed mathematical methods which permit us to study the nonequilibrium nature of the changes in the allele frequency distribution [Maruyama and Fuerst, 1984, 1985], and how these changes cause loss of alleles. Calculations are performed using a diffusion model of the nonequilibrium situation which follows a sudden change from large to small population size. The resulting equations are solved using a method of stochastic integrals developed by Maruyama [1980]. Populations are assumed to have an initial large size, and to exhibit a corresponding level of genetic variability which reflects an equilibrium between mutation and drift. For the results presented here, we examined three levels of starting genic heterozygosity, $H = 0.05$, 0.10 and 0.20. Other less realistic levels of variation were studied for our other publications, which dealt principally with the development of the theory. The population size is changed suddenly from a large to a small size (either 10, 20, 50, 100, or 200, depending on the situation being studied) and the population remains constant at this small size. Some of our results are summarized in Figures 2 and 3, which allow a comparison of the effects of different starting levels of genetic variation and different sample sizes.

Figure 2 shows the decline in the number of alleles present in the population for different sample sizes for the first 20 generations of the process. Not unexpectedly, variation is lost more rapidly for smaller sample population sizes. With samples of size 200 (Figure 2e), relatively little allelic variation is lost following the initial sampling event. Figure 2 can be used to plan the management of most mammalian species, especially by examining the curves for $H = 0.05$.

Figure 3 shows the effect of different levels of heterozygosity on the loss of alleles. Initial allele loss is greater in the more heterozygous populations. Note the very large loss of alleles in the first generation. This primarily due to the loss of rare alleles during the initial sampling process. More variable (both in terms of alleles and heterozygosity) populations have larger numbers of rare alleles, and these alleles are easily lost from the population in the early stages of sampling and management.

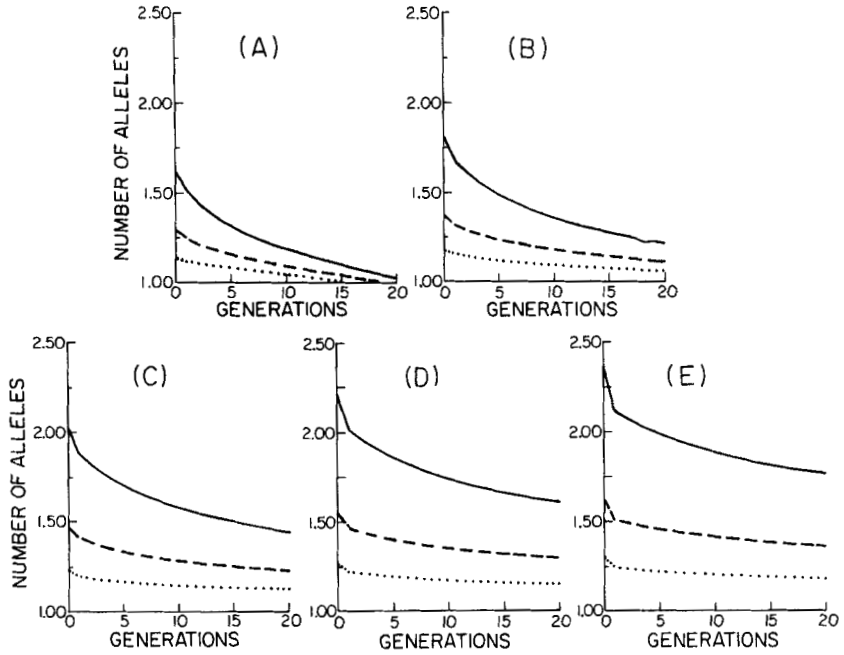

Fig. 2. The loss of the number of alleles (Na) in populations of different size (N). Three different starting levels of genic heterozygosity are shown for each case. (A) N = 10, (B) N= 20, (C) N = 50, (D) N = 100, (E) N = 200. ——, H = 0.2; — — —, H = 0.1; · · ·, H = 0.05.

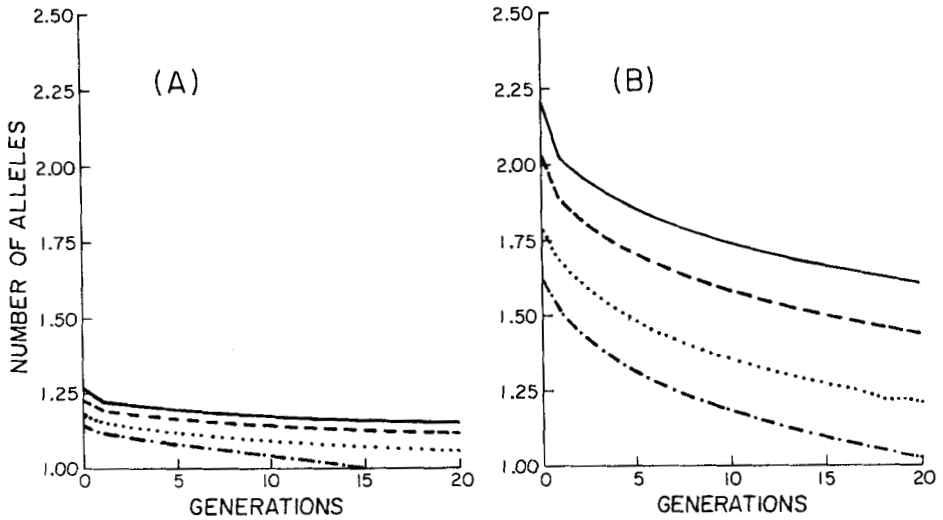

Fig. 3. Comparison of the effect of different starting average heterozygosity on the loss of the number of alleles (Na) in populations of different sizes. (A) starting heterozygosity = 0.05, (B) starting heterozygosity = 0.20. ——, N = 100; — — —, N = 50; · · ·, N = 20; —·—, N = 10.

Note that the distributions in Table 1 represent selectively equivalent alleles. If low frequency deleterious alleles are frequent, the rare allele frequency class will be even larger, and an even greater number of alleles will be lost in the initial generations. The fact of more rare alleles in the more variable populations is translated into a more rapid loss of alleles in the early generations, compared to less variable populations. Note the very marked drop in number of alleles in the very first generation in the population beginning with 0.20 average heterozygosity and a sample size of 200 (Figure 2e). In contrast, the populations which are shown in Figure 3a, all of which begin with average heterozygosity of 0.05, show a slow gradual decline in allele numbers with proportionate loss of alleles much lower than that seen in Figure 3b. It is important to keep in mind that most large mammal populations have been observed to have heterozygosities below 0.05.

A consideration of the allele frequency distributions from Table 1 indicates that several factors contribute to the pattern of loss of allelic variability in populations seen in Figure 2. First, for any heterozygosity level in the initial population, many alleles are rare in the population, and these alleles will almost all be lost in the sampling that occurs in the generations following the initiation of a small managed population. The rare frequency alleles represent a large proportion of the alleles encountered. They are also more abundant in the populations with higher initial levels of heterozygosity (Table 1). The result of this rapid loss of rare alleles which contribute little to heterozygosity in the population is that allele number declines much more rapidly than does heterozygosity during the early stages of population restriction. The same results, given in more detail, were found in our more extensive studies [Maruyama and Fuerst, 1985]. In our theoretical studies, this is a result of the assumption that there is a direct relationship between population size and genetic variability. This is not just a theoretical finding, however. Rare alleles are more frequent in more variable populations which have been described using electrophoretic techniques [see Chakraborty et al, 1980, for a further discussion of the pattern of allele frequency distributions in natural populations of both vertebrates and invertebrates].

Denniston [1978] and Allendorf [1986] indicated correctly that alleles at low frequency are lost more rapidly from the population. Following the initial loss of alleles due to the sampling of the founder population, there follows a less rapid loss of alleles due to the restricted size of the population. Alleles with intermediate frequencies tend to be retained in the population for a long period of time, while monomorphic alleles must await new mutations before they can be replaced. Both classes will probably be retained for the history of a managed population (unless the population size is very small). It must be kept in mind that we are no longer dealing with a population which is in equilibrium between mutation and random drift. Because of the small sizes of most captive populations, we can essentially ignore the effects of mutations for a period of time. However, it is clear that some mutation can occur, and the effect of such mutations will be to restore much of the variance for quantitative characters.

CONCLUSIONS

There are several implications of the results presented above. The sampling pattern which depends on the frequency distributions of both allele frequencies and

single locus heterozygosities indicates that ignoring allelic diversity when making conservation decisions is likely to produce a set of populations with very similar, or identical, allelic constitutions. These populations will contain a few highly variable genetic loci, coupled with little other allelic variation. Nevertheless, the average heterozygosity may be relatively high. This value would be misleading, however, since most allelic variants will have been lost. We have shown here and elsewhere [Maruyama and Fuerst, 1985] that allelic variation is lost much more quickly than genic variation (heterozygosity), and can only be restored by waiting for new mutations to occur. Under such a situation, a conservation program must carefully consider decisions which ignore alleles in favor of heterozygosity. Even when economic considerations make it easier to manage based on heterozygosity alone, we must strive to increase sample sizes, since an allele lost now cannot be recovered.

ACKNOWLEDGMENTS

This work was partially supported by a grant from the National Science Foundation, and from the Japanese Ministry of Education Science and Culture. We thank Jon Ballou and Kathy Ralls for their patience during our preparation of this manuscript.

REFERENCES

Allendorf, F. Genetic drift and the loss of alleles versus heterogosity. ZOO BIOLOGY 5:181–190, 1986.

Beardmore, J.A. Extinction, survival and genetic variation, pp. 125–151 in GENETICS AND CONSERVATION: A REFERENCE FOR MANAGING WILD ANIMAL AND PLANT POPULATIONS. C.M. Schonewald-Cox, S.M. Chambers, B. MacBryde, L. Thomas, eds. Menlo Park, CA, Benjamin/Cummings, 1983.

Chakraborty, R.; Fuerst, P.A.; Nei, M. Statistical studies of protein polymorphism in natural population. III. Allele frequency distributions. GENETICS 94: 1029–1063, 1980.

Chapman, C.G.D. On the size of a genebank and the genetic variation it contains, pp. 102–119 in CROP GENETIC RESOURCES: CONSERVATION & EVALUATION. J.H.W. Holden, J.T. Williams, eds. Boston, George Allen & Unwin, 1984.

Denniston, C. Small population size and genetic diversity. Implications for endangered species, pp. 281–290 in ENDANGERED BIRDS: MANAGEMENT TECHNIQUES FOR PRESERVING ENDANGERED SPECIES, S.A. Temple, ed. Madison, University of Wisconsin Press, 1978.

Fuerst, P.A.; Chakraborty, R; Nei, M. Statistical studies of protein polymorphism in natural populations. I. Distribution of single locus heterozygosity. GENETICS 86:455–483, 1977.

Huettel, M.; Fuerst, P.A.; Maruyama, T.; Chakraborty, R. Genetic effects of multiple bottlenecks in the Mediterranean fruit fly (*Ceratitis capitata*). GENETICS 94:s47–s48, 1985.

Kimura, M. Stochastic processes and the distribution of gene frequencies under natural selection. COLD SPRING HARBOR SYMPOSIUM QUANTITATIVE BIOLOGY 20:33–53, 1955.

Kimura, M.; Crow, J.F. The number of alleles that can be maintained in a finite population. GENETICS 49:725–738, 1964.

Maruyama, T. Stochastic Intergrals and their application to molecular evolution, pp. 151–166 in MOLECULAR EVOLUTION, PROTEIN POLYMORPHISM AND THE NEUTRAL THEORY. M. Kimura, ed. Berlin, Springer-Verlag, 1980.

Maruyama, T; Fuerst, P.A. Population bottlenecks and nonequilibrium models in population genetics. I. Allele numbers when populations evolve from zero variability. GENETICS 108: 745–763, 1984.

Maruyama, T.; Fuerst, P.A. Population bottlenecks and nonequilibrium models in population genetics. II. Number of alleles in a small population that was formed by means by a recent bottleneck. GENETICS 111:675–689, 1985.

Ryman, N.; Reuterwall, C.; Nygren, K.; Nygren, T. Genetic variation in Scandinavian Moose (*Alces alces*): Are large mammals monomorphic? EVOLUTION 34:1037–1049, 1980.

Wooten, M.C.; Smith, M.H. Large mammals are genetically less variable? EVOLUTION 39: 210–212, 1985.