11-1-2001

Determinants of Effective Population Size for Loci with Different Modes of Inheritance

Jay F. Storz
University of Nebraska - Lincoln, jstor2@unl.edu

U. Ramakrishnan
University of California at San Diego

S. C. Alberts
Duke University

Follow this and additional works at: http://digitalcommons.unl.edu/bioscistorz
Part of the Genetics and Genomics Commons

http://digitalcommons.unl.edu/bioscistorz/35

This Article is brought to you for free and open access by the Papers in the Biological Sciences at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Jay F. Storz Publications by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.
Determinants of Effective Population Size for Loci with Different Modes of Inheritance

J. F. Storz, Department of Ecology and Evolutionary Biology, University of Arizona, Tucson, AZ 85721
U. Ramakrishnan, Department of Biology, University of California at San Diego, La Jolla, CA 92093
S. C. Alberts, Department of Biology, Duke University, Durham, NC 27708

Corresponding author — Jay F. Storz

Abstract

Here we report an assessment of the determinants of effective population size (\(N_e\)) in species with overlapping generations. Specifically, we used a stochastic demographic model to investigate the influence of different life-history variables on \(N_e/N\) (where \(N = \) population census number) and the influence of sex differences in life-history variables on \(N_e\) for loci with different modes of inheritance. We applied an individual-based modeling approach to two datasets: one from a natural population of savannah baboons (Papio cynocephalus) in the Amboseli basin of southern Kenya and one from a human tribal population (the Gainj of Papua New Guinea). Simulation-based estimates of \(N_e/N\) averaged 0.329 for the Amboseli baboon population (SD = 0.116, 95% CI = 0.172 – 0.537) and 0.786 for the Gainj (SD = 0.184, 95% CI = 0.498 – 1.115). Although variance in male fitness had a substantial impact on \(N_e/N\) in each of the two primate populations, ratios of \(N_e\) values for autosomal and sex-linked loci exhibited no significant departures from Poisson-expected values. In each case, similarities in sex-specific \(N_e\) values were attributable to the unexpectedly high variance in female fitness. Variance in male fitness resulted primarily from age-dependent variance in reproductive success, whereas variance in female fitness resulted primarily from stochastic variance in survival during the reproductive phase.

The variance effective size of a population (\(N_e\)) can be estimated by measuring demographic parameters that influence the rate of increase in the variance of allelic frequencies (Caballero 1994; Crow and Denniston 1988). Such estimates provide a view of the prospective rate of genetic drift and permit inferences about the current trajectory of microevolutionary change. Effective population size is a parameter of central importance in population genetics and there is a great need for empirical assessments of the demographic determinants of \(N_e\) in natural populations. Such assessments are needed to determine whether \(N_e\) varies in a predictable fashion according to changes in life-history characteristics within and among species (Frankham 1995; Lande and Barrowclough 1987; Nei and Tajima 1981; Waite and Parker 1996). Estimates of \(N_e\) are also of value in testing ideas about the importance of genetic drift relative to other evolutionary forces in species with different modes of population structure (Lande 1979; Wright 1977).

In populations of constant size, \(N_e\) is primarily determined by variance in lifetime reproductive success (Wright 1931, 1938). In age-structured populations, the ratio of \(N_e\) to adult census number (\(N\)) is expected to fall within the range 0.25–0.75 under most demographic circumstances (Nunney 1993, 1996; Nunney and Elam 1994; Waite and Parker 1996). It has often been assumed that \(N_e/N\) for human populations is roughly 1/3 (where the census number \(N\) includes all age classes; Jorde 1980). However, this generalization is based on classical theory (Wright 1969) that does not factor in the effects of overlapping generations.

Sex differences in life-history variables also have implications for relative levels of genetic diversity at loci with different modes of inheritance. Under the infinite-sites model of DNA sequence evolution, the equilibrium level of neutral variation is determined by the product of \(N_e\) and the mutation rate (Kimura 1971). Thus, in the absence of sex differences in mutation rate, relative levels of diversity at loci with different modes of inheritance are solely determined by relative \(N_e\) values. It is common practice to calculate expected levels of diversity at such loci using discrete-generation formulas that assume a Poisson distribution of net lifetime progeny production for both sexes.
Figure 1. Influence of the operational sex ratio (Nf/Nm) on effective population size for loci with different modes of inheritance. Ne and Nf are the number of breeding males and females, respectively. Effective sizes for autosomal, X-linked, and mitochondrial loci (NeA, NeX, and NeMit, respectively) were calculated using equations 6.7, 6.8, and 6.15a in Hedrick (2000) with Nf = (10,000 - Nm). Calculations were based on the following assumptions: constant population size, nonoverlapping generations, and Poisson variance in fitness. The vertical dashed line indicates the point at which the curve for NeA intersects the curves for NeX and NeMit (at Nf/Nm = 7), indicating that locus-specific Ne values are equivalent at that point. The results are independent of the total population size (= Nf + Nm) used in this particular example.

(Wright 1969). Accordingly, Ne for Y-linked and mitochondrial loci (NeA and NeMit, respectively) are expected to be one-fourth Ne for autosomal loci (NeA). The expected fourfold reduction in Ne results from a twofold reduction due to haploidy and a further twofold reduction due to uniparental transmission (patrilineral transmission for the Y-chromosome, and matrilineral transmission for mtDNA). Similarly, Ne for X-linked loci (NeX) is expected to be three-fourths NeA due to hemizygosity in males (assuming that males are the heterogametic sex). However, these expectations may not hold up in the presence of non-Poisson variance in reproductive success. This non-Poisson variance can include intersexual variance (due to departure from a 1:1 breeding sex ratio) as well as intrasexual variance (due to the mating system).

For example, in polygynous species characterized by a high variance in male reproductive success, Ne for males is likely to be substantially lower than that for females. Under such circumstances, NeX/NeA will be greater than 0.75, NeM/NeA will be greater than 0.25, and NeMit/NeA will be less than 0.25. As illustrated in Figure 1, relative values of NeA, NeX, and NeMit change as a function of the operational sex ratio (Nf/Nm). Relative Ne values for loci with different modes of inheritance conform to expected values only when Nf/Nm is at unity. When Nf/Nm > 7, NeX and NeMit will both exceed NeA (and NeMit will exceed NeX; Figure 1). Thus, relative Ne values for loci with different modes of inheritance may be expected to deviate from Poisson-expected values in species characterized by polygynous mating systems and/or skewed sex ratios.

In populations with overlapping generations, age-dependent survival and fertility may provide an additional source of non-Poisson variance in net reproductive success (Charlesworth 1994; Hill 1972; Nunney 1993, 1996). Depending on the life-history characteristics of the species in question, sex differences in age-specific rates of survival and fertility may either offset or exacerbate the disparity in Ne values for loci with different modes of inheritance (Charlesworth 2001). In human population genetics, comparative analyses of loci with different modes of inheritance have figured prominently in studies of demographic history (Harpending and Rogers 2000; Hey 1997; Mesa et al. 2000; Wall and Przeworski 2000) and in studies investigating the role of selection and recombination in shaping patterns of DNA sequence variation (Nachman 2001; Nachman et al. 1998). It is therefore important to formulate sound predictions regarding the equilibrium level of neutral variation at such loci.

Estimating Effective Sizes of Natural Populations

Wright (1931, 1938) formulated the basic theoretical framework that described how Ne is affected by fluctuating population size, breeding sex ratio, and intrasexual variance in reproductive success. Hill (1972, 1979) derived the relationship that generalized Wright’s results to age-structured populations under conditions of demographic equilibrium (see also Emigh and Pollak 1979; Felsenstein 1971; Johnson 1977; Pollak 1980). The model contains one set of parameters related to demographic recruitment and another set related to variance in reproductive success. Application of Hill’s (1972, 1979) model to natural populations has been hindered by the difficulty of estimating parameters such as the variance and covariance of lifetime reproductive success among adults of both sexes (Harris and Allendorf 1989; Wood 1987). For long-lived animal species such as primates, only long-term longitudinal studies can provide sufficiently detailed life-history data to estimate the requisite parameters.

Provided that sufficiently detailed life-history data are available (i.e., age-specific rates of survival and fertility), Ne/N can be estimated using an individual-based simulation model (Rockwell and Barrowclough 1995). In contrast to traditional analytical methods, the individual-based simulation approach can be used to obtain confidence intervals for estimates of Ne/N and it provides a means of assessing the effects of demographic stochasticity. Here we report an empirical assessment of the determinants of Ne in age-structured populations. We applied an individual-based modeling approach to two datasets: one from a natural population of savannah baboons (Papio cynocephalus) in the Amboseli basin of southern Kenya and one from a human tribal population (the Gainj of Papua New Guinea). The purpose of this study was to investigate the influence of different life-history variables on Ne/N and the influence of sex differences in life-history variables on Ne for loci with different modes of inheritance.

Methods

Each of the primate datasets contained estimates of vital rates (age-specific rates of survival and fertility) for both sexes. Data for the Amboseli baboon population were taken from Alberts and Altman (2001) and data for the Gainj were taken from Wood (1987), Wood et al. (1985), and Wood and Smouse (1982). Since primates are the focus of this analysis, we are only considering transmission genetics relevant to mammals. However, results for X-linked loci are also generally applica-
able to any haplodiploid genetic system (Hedrick and Parker 1997). Likewise, results for Y-linked and mitochondrial loci are generally applicable to sex chromosomes and organelle genomes with uniparental inheritance.

Estimation of Lifetime Reproductive Success and $N_e$

To estimate $N_e$ for each of the two populations, we used an individual-based simulation model to calculate the variance in lifetime reproductive success of both males and females. Each simulation tracked the reproductive fate of each member of a complete age cohort over the course of a complete life cycle (i.e., newborn to newborn). The number of newborns produced by each member of a cohort of newborns was simulated using the empirical probability distributions of age-specific fertility and survival for both sexes. In each case the size of the simulated population was set equal to the observed census number ($N = 480$ for the Amboseli baboons and $N = 1318$ for the Gainj).

Simulations were initiated under conditions of demographic equilibrium. Survivorship curves for both sexes were used to estimate the stable age distribution, which was then used as a basis for sex and age assignments in the simulations. All newborns in the initial cohort were recorded as members of the same generation, and the numbers of male and female newborns were recorded as census numbers for each sex. Survival and reproduction were assessed over a period equal to the maximum life span recorded for each sex. The probability of reproduction by a given female in each year of that female’s life span was set equal to the age-specific fertility rate. In a similar manner, apportionment of the resultant offspring among males was determined by the normalized probability distribution of age-specific male fertility. Individual identities of newborns were recorded as part of the individual profiles of their parents. For both males and females, the length of each individual’s life span was determined by the age- and sex-specific probability distributions of survival. Deaths were recorded to assess lifetime reproductive success at the end of each simulation. Annual progeny production was summed over the total number of years comprising each individual’s reproductive life span to obtain the net lifetime production of male and female progeny by parents of each sex. For the Gainj, Wood’s (1987:177) matrilineal data were used to determine the number of wives per male as a function of male age. Each simulation stored information on marriage partners for the purpose of apportioning offspring among specific couples.

For each of the two primate populations, $N_e$ was estimated using the Hill (1972, 1979) equation:

$$
\frac{1}{N_e} = \frac{1}{16N_a} \left[ 2 + \text{var}(X_{mm}) \right]
+ \frac{2}{N_a} \left( \frac{N_a}{N_i} \right) \text{cov}(X_{mm}, X_{m})
+ \left( \frac{N_a}{N_i} \right)^2 \text{var}(X_{m})
+ 1 \left( \frac{1}{16N_a} \right) \left[ 2 + \text{var}(X_i) \right]
+ \frac{2}{N_a} \left( \frac{N_a}{N_i} \right) \text{cov}(X_{m}, X_{m})
+ \left( \frac{N_a}{N_i} \right)^2 \text{var}(X_{m})
\right]^{-1},
$$

where $N_i$ is the number of i-sexed individuals entering the breeding population each generation and $X_n$ is the net lifetime production of j-sexed progeny by i-sexed parents (with $m$ and $f$ denoting males and females, respectively). Simulations were repeated 2000 times to investigate the effects of demographic stochasticity on $N_e$. For each population the same number of simulations were repeated using a uniform distribution of male fertility (i.e., fertility of each age class was set equal to the overall mean). This permitted an evaluation of how $N_e$ was influenced by age-dependent variance in male reproductive success. In order to facilitate comparisons between species, we report estimates of the ratio $N_{ei}/N_i$, where $N$ is the census number of individuals of all ages. We also report estimates of $N_{ei}/N_{i(\text{adults})}$, where $N_{i(\text{adults})}$ is the census number of breeding adults (the only census number considered in previous studies; Nunney and Elam 1994).

To calculate $N_e$ for X-linked loci, equation 1 was modified to account for the fact that two-thirds of X-linked alleles are carried by females (and transmitted to male and female progeny), whereas one-third are carried by males (and transmitted to female progeny only; Pollak 1990):

$$
\frac{1}{N_{ex}} = \frac{1}{9N_m} \left[ 1 + \left( \frac{N_e}{N_i} \right)^2 \text{var}(X_{ex}) \right]
+ \frac{2}{N_m} \left( \frac{N_m}{N_i} \right) \text{cov}(X_{ex}, X_{m})
+ \left( \frac{N_m}{N_i} \right)^2 \text{var}(X_{m})
\right]^{-1},
$$

Effective sizes for Y-linked and mitochondrial loci were calculated as

$$
\frac{1}{N_{ef}} = \frac{2 + 2 \text{var}(X_{em})}{N_m}
\quad \text{and}
\frac{1}{N_{efl}} = \frac{2 + 2 \text{var}(X_{mf})}{N_i}
$$

respectively.

### Results

Estimates of vital rates for the Amboseli baboons and the Gainj indicated that both populations were approximately stationary in size and age composition (see Alberts and Altman 2001; Wood 1987). A total of 2000 Monte Carlo simulation trials were used to generate distributions of sex-specific estimates of variance in lifetime reproductive success (Table 1). In the Amboseli baboon population, variance in fitness was much higher among males than among females due to the high age dependence of male mating success (Alberts and Altman 2001). The resultant estimates of $N_{ei}/N_i$ averaged 0.329 (SD = 0.116, 95% CI = 0.172-0.537; Table 1 and Figure 2). When age-specific fertility of males

<table>
<thead>
<tr>
<th>Species</th>
<th>$N_{ei}/N_{i(\text{adults})}$ Mean (SD)</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P.\ cynocephalus$ (Amboseli)</td>
<td>0.508 (0.185)</td>
<td>0.274-0.854</td>
</tr>
<tr>
<td>$H.\ sapiens$ (Gainj)</td>
<td>1.270 (0.342)</td>
<td>0.873-1.640</td>
</tr>
</tbody>
</table>

For each population, estimates of the mean, standard deviation, and 95% confidence interval were based on 2000 Monte Carlo trials.
was held constant, this mean value increased to 0.360 (SD = 0.299, 95% CI = 0.211–0.564; Figure 2). Thus, \( N_e/N_t \) was reduced by 9% as a result of age-dependent variance in male reproductive success. However, sex differences in life-history variables were not sufficient to cause \( N_{ea}/N_{ea}^* \), \( N_{ev}/N_{ea}^* \), or \( N_{emo}/N_{ea}^* \) to deviate significantly from Poisson-expected values (Table 2).

Table 2. Simulation-based estimates of locus-specific \( N_e \) ratios for two primate populations

<table>
<thead>
<tr>
<th>Population</th>
<th>( N_e/N_m ) (exp. = 0.75)</th>
<th>( N_{ea}/N_{ea}^* ) (exp. = 0.25)</th>
<th>( N_{ev}/N_{ea}^* ) (exp. = 0.25)</th>
<th>( N_{emo}/N_{ea}^* ) (exp. = 0.25)</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>R. cynocephalus</em> (Amboseli)</td>
<td>1.065 ± 0.329 (0.618–1.667)</td>
<td>0.740 ± 0.094 (0.598–0.897)</td>
<td>0.229 ± 0.088 (0.129–0.388)</td>
<td>0.289 ± 0.120 (0.138–0.520)</td>
</tr>
<tr>
<td><em>H. sapiens</em> (Gainj)</td>
<td>0.984 ± 0.184 (0.714–1.315)</td>
<td>0.690 ± 0.057 (0.605–0.791)</td>
<td>0.211 ± 0.056 (0.141–0.316)</td>
<td>0.218 ± 0.061 (0.133–0.326)</td>
</tr>
</tbody>
</table>

\( N_e/N_m \) is the operational sex ratio (female/male), and \( N_{ea}, N_{ev}, N_{emo} \) are the effective population sizes for autosomal, \( X \)-linked, \( Y \)-linked, and mitochondrial loci, respectively. For each population, estimates of the mean (±1 SD) and the 95% confidence interval (in parentheses) were based on 2000 Monte Carlo trials.

In the Gainj population, variance in reproductive success was similar for both sexes. All else being equal, the fact that some men had multiple wives could be expected to increase the variance in male reproductive success relative to that of females (Wood 1987). However, variance in lifetime reproductive success of both sexes was largely attributable to the stochastic nature of survival during the reproductive phase. Because the reproductive life spans of females were, on average, longer than those of males (due to a younger age at marriage), this source of variance in female reproductive success was sufficient to offset the increased variance in male reproductive success that resulted from polygynous mating. Simulation-based estimates of \( N_e/N_t \) for the Gainj averaged 0.786 (SD = 0.184, 95% CI = 0.498–1.115; Table 1 and Figure 3). When age-specific fertility of males was held constant, this mean value increased to 1.107 (SD = 0.299, 95% CI = 0.611–1.620; Figure 3). Thus \( N_e/N_t \) was reduced by 41% as a result of age-dependent variance in male reproductive success. In the Gainj, \( N_{ea}/N_{ea}^* \) was reduced below the expected value of 0.75, and both \( N_{ev}/N_{ea}^* \) and \( N_{emo}/N_{ea}^* \) were decreased below the expected value of 0.25. However, in each case the 95% confidence interval bracketed the Poisson-expected value (Table 2).

Discussion

The estimated ratio of \( N_e \) to adult census number (\( N_{eadult} \)) for the Amboseli baboon population (0.508) was intermediate between estimates obtained for other polygynous species with overlapping generations, such as the short-nosed fruit bat, *Cynopterus sphinx* (\( N_e/N_{eadult} = 0.42; \) Storz et al. 2001), and the grizzly bear, *Ursus arctos* (\( N_e/N_{eadult} = 0.82; \) Nunney and Elam 1994). By contrast, \( N_e/N_{eadult} \) ratio for the Gainj population (1.270) greatly exceeded the range of values (0.25–0.75) that are generally expected for age-structured populations under normal demographic conditions (Nunney 1993, 1996). Moreover, our average simulation-based estimate of \( N_e/N_t \) exceeded the estimate obtained by Wood (1987:181) in his original analysis of the Gainj data (0.786 versus 0.494). There are several explanations for this discrepancy.

In our individual-based simulation model, lifetime reproductive success was measured as the number of newborns produced by each member of a cohort of newborns. By contrast, Wood’s (1987) estimate of variance in lifetime reproductive success considered only those individuals who survived to adulthood. It does not necessarily matter which stage of the life cycle is used as a reference point for measuring lifetime reproductive success so long as the denominator of the ratio \( N_e/N_t \) is chosen accordingly (Nunney and Elam 1994; Rockwell and Barrowclough 1995).
ting $N$ equal to the number of adults recruited into the breeding population, Nunney and Elam’s (1994:178) applied Wood’s (1987) data to a simplified version of the Hill equation and obtained a revised estimate of $N_e/N_{\text{adults}} = 1.27$ for the Gainj population. Although Nunney and Elam’s (1994) model was based on the assumption of age-independent survival, age-independent fertility, and invariant life-history schedules for both sexes, their point-estimate of $N_e/N_{\text{adults}}$ was identical to that obtained by our individual-based model (Table 1).

$N_e/N$ estimates were considerably higher for the Gainj relative to the Amboseli baboon population (mean = 0.786 versus 0.329). This is primarily due to the exceptionally long generation times of the Gainj (38.3 versus 10.1 years for males and 31.3 versus 10.2 years for females). In age-structured populations, $N_e$ is generally more sensitive to changes in parameters related to reproductive success (variances and covariances of net lifetime production of male and female progeny) than to changes in parameters related to recruitment (generation time and the number of males and females attaining sexual maturity each year; Harris and Allendorf 1989; Wood 1987). However, Wood (1987) suggested that recruitment-related parameters are more likely to vary according to local ecological conditions, and may therefore constitute the most important source of interpopulation variation in $N_e$. Compared with other human populations, the Gainj are characterized by unusually long generation times and low rates of recruitment (Wood et al. 1985; Wood and Smouse 1982). Thus, our estimate of $N_e/N$ for the Gainj is unlikely to be representative of human tribal populations in general. As stated by Nei (1970:695): “It is apparent that the $N_e/N$ ratio in human populations is never constant either spatially or temporally.”

The above discussion applies to the general case of $N_e$ for autosomal loci. Our estimates of relative $N_e$ values for loci with different modes of inheritance agree with Charlesworth’s (2001) conclusion that sex differences in life-history variables would have to be fairly extreme in order to produce significant departures from Poisson-expected values. Interestingly, locus-specific $N_e$ ratios closely conformed to Poisson-expected values not because variance in male fitness was low, but because variance in female fitness was unexpectedly high (Table 1). Variance in male fitness resulted primarily from age-dependent variance in reproductive success, whereas variance in female fitness resulted primarily from stochastic variance in survival during the reproductive phase.

It is important to note that locus-specific $N_e$ ratios may deviate from Poisson-expected values if male reproductive success is strongly influenced by age-independent phenotypic traits. Indeed, Charlesworth’s (2001) results indicate that age-independent variance in reproductive success (e.g., due to sexual selection) may have profound effects on sex-specific $N_e$ values.

In addition to the effects of sex-differences in life-history variables, relative levels of neutral genetic variation at loci with different modes of inheritance are determined by the effects of sex differences in rates of mutation and migration. With regard to neutral mutation rates in humans, available evidence suggests that $N_e$ values are generally more than that of females (Anagnostopoulos et al. 1999; Bohossian et al. 2000; Nachman and Crowell 2000). And with regard to migration patterns in humans, available evidence suggests that females are generally more likely than males to leave their birthplace prior to reproduction, at least in agricultural societies characterized by patrilineal inheritance of land holdings (Burton et al. 1996). Consequently mtDNA markers may generally exhibit lower levels of genetic differentiation relative to Y-linked and autosomal markers (Seielstad et al. 1998), although this depends on the geographic scale under consideration (Mesa et al. 2000; Seielstad 2000). The opposite pattern has been documented in macaques due to a strong degree of female philopatry (Melnick and Hoelzer 1992). In principle, the effects of sex-biased dispersal on patterns of

![Figure 3](image)

**Figure 3.** Simulation-based estimates of (A) the variance in male fitness and (B) $N_e/N$ for the Gainj tribe of Papua New Guinea. Distributions were generated from 2000 Monte Carlo trials using an individual-based model. Simulations based on the actual vital rates were used to generate the "observed fertility" values. Simulations based on uniform male fertility were used to generate the "uniform fertility" values (see text for details).
variation at sex-linked loci can be factored in using the “neighborhood adjustment” to $N_e$ (Maruyama 1972; Wright 1943). If samples are obtained from a geographically restricted portion of a species’ range, differences in levels of variation at sex-linked loci may often reflect sex differences in migration (and hence genetic neighborhood size).

Contrasting patterns of sex-biased dispersal are well illustrated by the primate populations chosen for analysis in this study. In the Gainj, prereproductive dispersal is highly female biased (Wood et al. 1985). As stated by Wood (1987:184), “Combining the effects of differential dispersal and differential reproductive heterogeneity, it could be claimed that the male effective population size is considerably smaller than the female effective size in the Gainj.” And with regard to comparative levels of variation at paternally and maternally transmitted genes, “it would be meaningful to predict that the extent of genetic drift is greater in males than in females.” By contrast, in the Amboseli baboon population, prereproductive dispersal is mediated almost exclusively by males (Alberts and Altmann 1995). In this case, sex-specific neighborhood sizes may be more closely similar since the higher reproductive variability in males may be offset by the fact that paternally transmitted genes are also more likely to be of immigrant origin.

References


