The effective population number estimated in humans is \( \sim 10,000 \) (Nei and Graur 1984; Nei 1987; Takahata 1993). This rather small number may contradict the fact that humans had been occupying vast areas in Africa and Eurasia for a long time since *Homo erectus* first migrated out of Africa \(-2\) Mya (e.g., see Lewin 1984; Fagan 1990). One possible cause is a recent bottleneck, or a sudden reduction in population size, but a severe bottleneck is incompatible with the unusual polymorphism at the major histocompatibility complex (*Mhc*) loci (Takahata 1990; Klein et al. 1993). In this letter, I argue that a more likely cause is frequent extinction and recolonization (turnover) of subpopulations in the lineage leading to modern humans.

Wright (1931) introduced the concept of effective population number \( N_e \) in order to reduce complicated breeding structure in actual populations to the equivalent ideal case (Crow and Kimura 1970, p. 345). In an ideal population, \( N_e \) is equal to the number of breeding individuals, whereas in an actual population \( N \) depends on various factors such as sex ratio, differences in offspring numbers, temporal changes in population size, and overlapping generations (for review, see Wright 1969). Wright (1969) also pointed out that if (1) a species has subdivided population structure and (2) local populations are liable to frequent extinction with restoration of a few stray immigrants, the species passes repeatedly through extremely reduced states of \( N_e \). Slatkin (1977), as well as Maruyama and Kimura (1980), formulated this problem on the basis of a traditional approach in population genetics: that is, to use the variance of gene frequencies and the probability of identity of two randomly chosen genes. Maruyama and Kimura (1980) derived formulas of \( N_e \) under the finite-island model of population structure (Malécot 1951; Maruyama 1970) and the stepping-stone model (Kimura and Weiss 1964). Their approximate formula for a diploid species in the finite-island model is given by

\[
N_e = N + \left( Nm + \frac{1}{4} \right) \frac{n}{\lambda + m},
\]

where \( N \) is the number of breeding individuals in each subpopulation, \( m \) is the migration rate, \( n \) is the number of subpopulations, \( N_e \) is the number of breeding individuals in each subpopulation (or the harmonic mean of \( N \) when it changes with time), and \( \lambda \) is the turnover rate per subpopulation. When \( \lambda \rightarrow 0 \), \( N_e \) cannot be \( < Nn \), whereas if \( \lambda \gg m \), \( N_e \) reduces to \( N \) or \( n/(4\lambda) \), whichever is greater. Nei and Takahata (1993) showed that the coalescence time of neutral genes is well predicted from the equivalent formula in a randomly mating population if we use appropriate \( N_e \). The coalescence approach appeals more to our intuition and helps us to better understand the underlying evolutionary process. However, Nei and Takahata did not consider the possibility of subpopulation turnovers. This is done as follows.

We consider ancestral lineages of two genes that are sampled either from the same subpopulation or from different subpopulations, under the finite-island model. The coalescence of genes can occur only when they or their ancestral lineages are in the same subpopulation. Without migration, the coalescence time of two genes that are sampled from the same subpopulation is exponentially distributed with mean \( 2N \) (for review, see Nei 1987). When gene migration is permitted, there are two possibilities: (a) coalescence before migration and (b) migration before coalescence. In the first case, the time \( (\tau) \) until event \( (a) \) is exponentially distributed:

\[
g(\tau) = \frac{1}{2N_e} \exp \left[ -\left( \frac{1}{2N_e} + 2m \right) \tau \right],
\]

where \( e^{-2mt} \) represents the probability of no gene migration for \( \tau \) generations. From formula (2), the probability of \( (a) \) is \( 1/(1+4Nm) \), and the mean of \( \tau \) is \( 2N/(1+4Nm)^2 \). Dividing the latter by the former, we have the mean time conditioned on \( (a) \): \( 2N/(1+4Nm) \). In the case of \( (b) \), gene migration occurs with probability \( 4Nm/(1+4Nm) \), and two ancestral lineages reside in different subpopulations. We denote by \( t_1 \) the mean coalescence time of two genes when they are sampled from different subpopulations. Slatkin (1991) expressed the mean coalescence time \( (t_0) \) of two genes sampled from the same subpopulation as

\[
t_0 = \frac{2N}{1+4Nm} + \frac{4Nm}{1+4Nm} t_1.
\]

On the other hand, when two genes are sampled from different subpopulations, the coalescence becomes possible only after the ancestral lineages happen to be in the same subpopulation. This may happen by
extinction of one subpopulation and recolonization by a duplicate of the other subpopulation. The analysis of Maruyama and Kimura (1980) suggests that for a particular pair of subpopulations, the time \( t \) until a subpopulation turnover is exponentially distributed:

\[
f(t) = \frac{2\lambda}{n-1} \exp\left( -\frac{2\lambda t}{n-1} \right),
\]

with mean \((n-1)/(2\lambda)\). It is noted that the process of subpopulation turnover is mathematically identical to that of allelic turnover under balancing selection (Takahata 1990). The differences between formula (4) here and formula (14) in Takahata (1990) come from different definitions of turnover rates. In addition to subpopulation turnovers, gene migration may occur. The probability distribution of the time at which both ancestral lineages come into the same subpopulation by migration is exponential with mean \(((n-1)/(2\lambda + 2m))\). Using this formula and \( t_0 \) in formula (3), we can express \( t_1 \) as

\[
t_1 = \frac{n-1}{2(\lambda + m)} + t_0.
\]

Since \( \lambda < 1 \) and \( m < 1 \), \( t_1 \) cannot be shorter than \((n-1)/4 \) generations. Formulas (3) and (6) lead to

\[
t_0 = 2N \left[ 1 + \frac{m(n-1)}{\lambda + m} \right].
\]

We note in formulas (6) and (7) that effects of subpopulation turnover and migration are indistinguishable in \( t_1 - t_0 \), but the former effect is more important than the latter in determining \( t_0 \). If \( \lambda = 0 \), \( t_0 = 2Nn \), but if \( \lambda \gg m \), \( t_0 = 2N \). It is this effect of subpopulation turnover that shortens the coalescence time substantially.

To derive \( N_e \), we must consider a random sample of two genes from the whole population (Slatkin 1991; Nei and Takahata 1993). The probability that two genes are sampled from the same subpopulation is \( 1/n \), and the probability that they are sampled from different subpopulations is \( 1 - 1/n \). Using formulas (6) and (7), we obtain the mean coalescence time \( T \) for such a sample of two genes:

\[
T = \frac{1}{n} t_0 + \frac{n-1}{n} t_1
= 2N + \left( 2Nm + \frac{n-1}{2n} \right) \frac{n-1}{\lambda + m}.
\]

If we equate formula (8) to \( 2N_e \) (the expected value in a randomly mating population with size \( N_e \)), we obtain

\[
N_e = N + \left( Nm + \frac{n-1}{4n} \right) \frac{n-1}{\lambda + m}.
\]

The difference between formula (9) here and Maruyama and Kimura’s formula is due to their approximation for large \( n \). In terms of mean coalescence times, we can also express Wright’s (1931) \( F_{ST} \) as \( 1 - t_0/T \) (Slatkin 1991):

\[
F_{ST} = \frac{1}{1 + 4Nn(h + mn)/(n-1)^2}.
\]

If \( n \gg p = \lambda/m, F_{ST} \) is well approximated by \( 1/(1 + 4Nm) \), as shown by Wright (1931).

We can use formulas (9) and (10) to make inferences on the demographic history of humans over the past 1 Myr. It is believed that until the Agricultural Revolution, humans lived in groups of \(~25\) individuals and that the (census) population size was \(~10^7\) (e.g., see Lewin 1984, p. 94). We may assume that the total number of breeding individuals \((Nn)\) was \(~10^6\), much greater than the effective number \((N_e)\) estimated from genetic data. It is not obvious that the breeding unit was a single group of \(~25\) individuals; in our context, \( N \) should be referred to as the number of breeding individuals in a unit within which they can mate more or less at random. A question is how large \( N \) was, or, equivalently, how many such breeding units \((n)\) existed in the human population. We consider this question under the condition of \( Nn \gg N_e > N \). This implies that \( 100 \ll n < 10,000 \), when \( Nn = 10^6 \) and \( N_e = 10,000 \).

If \( p = \lambda/m < 1, N_e > Nn \) must hold in formula (9). However, since this is incompatible with the above condition, the opposite case of \( \lambda > m \) (or \( p > 1 \)) is required. Formulas (1) and (9) also indicate that if \( 4Nnm > 1, p \sim 100 \), whereas if \( 4Nnm < 1, 4N \lambda \sim 100 \) (so that \( p > 100 \)). On the other hand, the average value of \( F_{ST} \) is \( 10\% \) at various loci (Nei 1987; Nei and Livshits 1990). A smaller \( F_{ST} \) is estimated if we use the data of allele frequencies at five human leukocyte antigen (HLA) loci, based on the report in the 11th Histocompatibility Complex Workshop (Y. Satta, personal communication). The \( F_{ST} \) value for three major ethnic groups is \(<3\% \) at all five loci. There is little local differentiation in the human population. For \( F_{ST} \sim 10\% \), it is necessary that \( 4N\lambda/n + 4Nnm \sim 9 \) in formula (10). This relationship cannot be met if \( 4Nnm < 1 \): since \( 4N \lambda > 100 \), both \( 4N\lambda/n \) with \( n \gg 100 \) and \( 4Nnm \) are \(<1 \). We can therefore conclude that \( 4Nnm \sim 9 \). In other words, the small extent of local differentiation in the human population is largely attributed to gene migration rather than to frequent extinction and recolonization of subpopulations.

We suppose that \( Nnm \sim 2.5 \) and \( N\lambda \sim 250 \), from which we draw two conclusions. First, \( N \) must be much greater than the number of \(~25\) individuals in a primitive group, and the breeding unit must have consisted of a number of such groups. For instance, if \( \lambda = 0.1 \), then \( N \sim 2,500 \) and \( n \sim 400 \). Second, the required turnover rate must be high. Even when \( N \) is nearly \( 10,000 \), each subpopulation survives for only \( 40 \) (=1/\( \lambda \)) generations, on average.

Frequent subpopulation turnover reduces \( N_e \) substantially, but there are other possibilities. One possibility
often invoked to explain the reduced $N_e$ is the effect of hitchhiking by advantageous mutations. Advantageous mutations sweep out polymorphism at linked loci, resulting in spuriously small $N_e$. In relation to subpopulation turnover, another possibility is that $N$ differs from one subpopulation to another. In fact, available genetic data have indicated more variation in Africans, suggesting large subpopulation numbers (e.g., see Cann et al. 1987; Nei 1987). The point is that the turnover rate may well be inversely correlated with $N$: The smaller $N$, the more liable to extinction. When *Homo sapiens* spread over the world in the late Pleistocene (e.g., see Lewin 1984; Fagan 1990), local populations in Eurasia might have undergone more frequent extinction than did those in Africa, because of the then-adverse environment. A notable example is the extinction of the Neanderthals. Perhaps, many more failures preceded the eventual success. Africans could then serve directly or indirectly as founding populations in Eurasia. These inferences argue for the Saharan pump (Fagan 1990, p. 65) and the out-of-Africa hypothesis.

Acknowledgments

This paper was supported in part by grant 05044127 in the Monbusho International Scientific Research Program: Joint Research, the Ministry of Education, Science and Culture, Japan.

LITERATURE CITED


JAN KLEIN, reviewing editor

Received March 1, 1994
Accepted April 29, 1994