The Founder Effect and Deleterious Genes

FRANK B. LIVINGSTONE

Department of Anthropology, University of Michigan

ABSTRACT   During the rapid growth of a population from a few founders, a single deleterious gene in a founder can attain an appreciable frequency in later generations. A computer simulation, which has the population double itself in early generations, indicates a lethal could attain a frequency of 0.1. Since deleterious recessive genes are eliminated from large populations at a very slow rate, variations in their frequencies in present major human populations may be due to the founder effect during earlier rapid expansion.

Many distinctive human populations are characterized by the presence of one or more lethal or severely deleterious genes in frequencies which would be defined as polymorphic according to Ford's ('40) famous definition. The particular genetic disorder, however, varies. The Old Order Amish of Lancaster County, Pennsylvania have a gene frequency of 0.07 for the recessive Ellis-van Creveld syndrome, while the Amish as a whole have a frequency of about 0.05 of the recessive cartilage-hair hypoplasia syndrome (McKusick et al., '64). Many of the tri-racial isolates of Eastern United States also have a high frequency of a deleterious gene (Witkop et al., '66). Although such populations are frequently defined by religious or ethnic criteria, there are others not so defined. Several island populations in the Åland archipelago have a gene frequency of greater than 0.1 for von Willebrand's disease (Eriksson, '61), and the Boer population of South Africa and some populations of Northern Sweden have frequencies of porphyria much greater than those of other populations (Dean, '63; Waldenstrom and Haeger-Aronsen, '67). However, these conditions are dominant and do not have the very severe effects of other hereditary disorders found in high frequencies. On the other hand the population of the Chicoutimi District of Quebec has recently been found to have a gene frequency of about 0.02 for tyrosinemia, which is a lethal recessive (Laberge and Dallaire, '67).

In most of these cases the population in question has undergone a rapid increase in recent years, and the question arises as to whether this rapid expansion and the original small size of the isolate could account for the high frequency of the deleterious gene. Such an explanation by the founder effect seems obviously to apply to most of the cases cited above, but the founder effect may well be a more general explanation of human gene frequency differences. It is now becoming apparent that the major populations of mankind vary significantly in their frequencies of deleterious genes and that many large populations such as Eastern European Jews have high frequencies of deleterious genes which are found in low frequencies in other populations (McKusick, '66). There have been many attempts to determine how such genes could be polymorphic, for example, Anderson et al. ('67) and Knudson et al. ('67) have discussed cystic fibrosis and Myrianthopoulos and Aronson ('66), Tay-Sachs disease. The purpose of this paper is to attempt to determine the extent to which the founder effect can cause high frequencies of deleterious genes with various models of population expansion.

The occurrence which initiated this research is the gene for sickle cell hemoglobin in the Brandywine isolate of Southeast Maryland. At present the sickle cell gene frequency in this isolate is about 0.1 (Rucknagel, '64). The high frequencies of this gene in many parts of Africa, India, and the Middle East are now well-accepted as being due to a relative resistance of the sickle cell heterozygote to falciparum malaria. The high frequency in the Brandywine isolate may have a similar explanation, but the surrounding Negro population does not have such a high frequency. And although the endemicity of falciparum

malaria in Southeast Maryland in the last century is not known in any detail, it would not appear to have been great enough to explain the high sickle cell frequency in the Brandywine isolate. The isolate also has many other deleterious genes in high frequency (Witkop et al., '66).

The Brandywine isolate seems to have had its beginning in the early Eighteenth Century when laws were passed to prohibit co-habitation and marriage among races, which prior to then were presumably frequent or at least known. Up to 1720 there were several prosecutions under these laws of individuals with surnames currently present in the isolate (Harte, '63). Harte ('63) has maintained that the Brandywine isolate is derived from these illegal unions, and Witkop et al. ('66) show that the most common surname came from such a union. In 1790 the first United States Census recorded 190 persons with the group's surnames as "other free people," and since then over 90% of the recorded marriages have been endogamous or between individuals with surnames within the group (Harte, '59). According to Harte ('59) there are six "core" surnames which have been associated with the group since its founding and comprise 66% of the population and another ten surnames which entered the group after the Civil War, but Witkop et al. ('66) list seven core surnames and eight marginal ones. The total population of the isolate is now estimated to be 5,128 (Witkop et al., '66), and the statistics do indicate rapid, if erratic, growth (Gilbert, '45; Harte, '63).

In order to simulate gene dynamics the population has been assumed to have doubled itself in early generations, and then after slower growth to have approached a doubling in recent generations. The simulation was run for 10 generations with the following numbers in succeeding generations: 20, 40, 80, 160, 320, 640, 664, 728, 856, 1112. This approximates the early demographic history of the Brandywine isolate, but the isolate is much larger today. However, gene frequency change in later generations with a large population is very small.

The simulation model randomly selects two parents from the initial population which has been assumed to have either one or two sickle cell heterozygotes among the founders. A family size is randomly determined, the offspring generated and then selected out or stored with no compensation for those not surviving (a copy of the program is available on request). Since the population is increasing rapidly, the family size distribution approximates that recorded by Roberts ('65) for a population in Tanzania which has about 4.0 surviving offspring per female. The founder population can actually vary in size, however. The size of the offspring generation is the number which is set; but with an average of 4.0 offspring per marriage and 40 offspring, the founder population would be expected to consist of ten marriages or 20 individuals.

Figure 1 shows the distribution of the deleterious gene frequencies after ten generations for two sets of 50 runs each with different initial conditions and different fitnesses for the genotypes. With a gene fre-
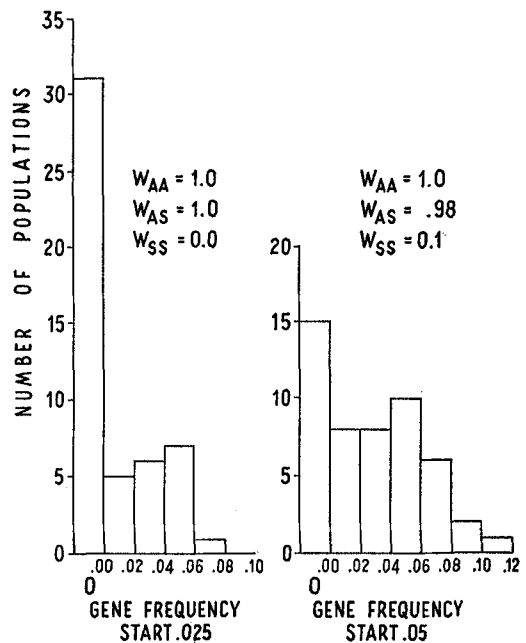

Fig. 1 The distributions of the frequency of the sickle cell gene (S) after ten generations of expansion for two sets of 50 runs each with different initial gene frequencies and different fitnesses for the genotypes (W's). Note: the populations in which the S gene has been completely eliminated are separated in the left column from those in which it is still present.

quency of 0.05, which is comparable to having two sickle cell heterozygotes among the founders, the gene is present at a frequency of greater than 0.04 in almost 40% of the populations, while for a starting gene frequency of 0.025 or one founder with the sickle cell trait, 16% of the populations have the gene at a frequency of greater than 0.04. With two founders there were runs which resulted in a gene frequency as high as that of the sickle cell gene in the Brandywine isolate, but with one founder there were none as high. However, there were many frequencies close to it, so that such an outcome is possible if not probable. Hence there seems to be no necessity to postulate a selective advantage for the sickle cell in the Brandywine isolate. It should be pointed out that this simulation and further ones assume the population is closed. Gene flow from other populations would tend to decrease the frequency of the deleterious gene, but if most of the population's expansion is due to natural increase, then the founder effect would be most important.

In order to determine whether such high frequencies could occur in a population with a greater number of founders, a similar program was run with 40 and 80 founders. The results are shown in figure 2. These runs were started with one carrier of the lethal gene and the population doubled itself for five generations, so that it ended with 1,280 and 2,560 individuals, respectively. Although the lethal gene is not present in high frequencies in as many populations, it is still present in about 5% in a frequency greater than 0.04. The fact that populations begun with a few founders should have such high frequencies of lethal genes seems to indicate that they can contribute to the problem of the genetic load. According to Morton ('60), the average individual has the equivalent of four recessive lethals in the heterozygous state. For a population with 40 founders this would imply over 100 lethal or deleterious genes among the founders, so that several would be expected to attain high frequencies. The fact that the number of deleterious genes in small populations started by a few founders seems to average much less may indicate a lower genetic load, but the Brandywine isolate,

the Amish, and the Eastern European Jews do have several deleterious genes in high frequencies. In any case, it seems to be a possible way to study the problem.

The most recent population expansion which seems to have increased the frequency of a lethal gene is the peopling of the Saguenay River and Lake St. John region by French Canadians. Settlement of the Upper Saguenay did not begin until the 1830's, and Chicoutimi was founded in 1840 by 220 individuals from La Malbaie, 66 from Eboulements, and 37 from Baie St. Paul (Buies, 1896). In 1861 the population of Chicoutimi was 10,478 and in 1871 it rose to 17,483. Much of this increase was undoubtedly due to immigration, but given the enormous rate of increase of the French Canadian population as a whole (Henripin, '54), the expansion of the population of Chicoutimi was due to a great extent to natural increase. Under these conditions a lethal gene frequency of 0.02 would not be unlikely and seems
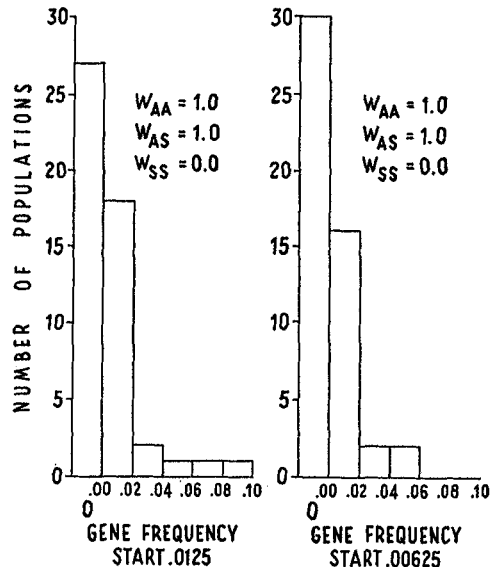

Fig. 2 The distributions of the frequency of a lethal recessive gene (S) after five generations of expansion for two sets of 50 runs each with different initial gene frequencies which represent one carrier of the lethal gene in a founder population of 40(0.0125) and one carrier in a population of 80(0.00625). Note: the populations in which the lethal gene has been completely eliminated are separated in the left column from those in which it is still present.

to agree with our simulation model. The fact that the entire French Canadian population stems from about 10,000 original settlers (Henripin, '54) may have led to this population having its own set of lethal genes.

The high frequencies of deleterious genes in the Eastern European Jewish populations of Lithuania and Eastern Poland may have a similar explanation, although there is disagreement about this possibility (McKusick, '66). Myrianthopoulos and Aronson ('66) do not consider such an explanation likely for the Tay-Sachs gene. Instead they propose a slight selective advantage for the heterozygote. They have to postulate the operation of this selective advantage for 50 generations which is longer than the population has been there. It is also much longer than the factor thought to confer the selective advantage, typhoid fever in the ghettos (Aronson, '64), seems to have been present as a serious disease. The Tay-Sachs gene attains its highest frequencies in the Jewish populations of Southern Lithuania and Northeast Poland, which were founded in the Twelfth Century after the Crusades led to the persecution of the Jews in Germany. Although the Jewish settlements in Lithuania were founded by refugees from the west, according to Herzog ('65) they preceded by two or three centuries the Jewish settlements in Mazovia to the west in Poland. Thus, these colonies were isolated for some time and were actually expanding to the west into Northern Poland when the Jews were expelled from Lithuania in 1495. Most moved to adjacent territories but then moved back to Lithuania in 1503. Hence, the population history of these Jewish groups seems to be one of expansion from a few founders. In any case by the time of the flowering of Eastern Jewish culture in the Sixteenth Century, the population was very large and continued to expand up to the Twentieth Century.

When a population size of 1,000 or more is attained, the change in gene frequency is approximated by deterministic equations. For a lethal recessive the frequency after n generations is: $q_n = \dfrac{q_0}{1 + nq_0}$,

where q_0 is the initial gene frequency. A lethal gene in a large population is thus

eliminated at a very slow rate, particularly when it occurs in a very low frequency. If the Tay-Sachs gene increased to 0.05 in the early generations before the population became large, in the approximately 30 generations since then, the gene would have decreased to $\dfrac{0.05}{1 + 30(0.05)} = 0.02$, which is about the frequency today in Eastern European Jews.

The fact that lethal genes are eliminated at such a slow rate in large populations would make it possible for them to have "polymorphic" frequencies long after the original expansion. Since most of the world's populations have expanded rapidly in the last 1,000 years, much of the variability in the frequencies of lethal genes (or non-lethals for that matter) could be a consequence of the original expansions of the major populations. As an example, it is suggested that this effect may explain the high frequencies of cystic fibrosis in the populations of Europe, which range around a gene frequency of 0.02. For the non-Caucasian populations on Hawaii Wright and Morton ('68) have estimated the gene frequency for cystic fibrosis to be 0.003, which presumably is close to the equilibrium frequency due to a balance of selection and mutation. Given this frequency, over 50% of a set of founder populations of size 100 would be expected to have a carrier of this lethal. With the sudden expansion of such a set of founder populations it seems possible that such a lethal could attain a frequency of 0.05 for the entire population. A more precise mathematical expression of the problem seems possible and could perhaps yield a solution.

LITERATURE CITED

Anderson, C. M., J. Allan and P. G. Johansen 1967 Comments on the possible existence and nature of a heterozygote advantage in cystic fibrosis. In: Cystic Fibrosis. E. Rossi and E. Stoll, eds. Bibliotheca Paediatrica, No. 86. S. Karger, New York, pp. 381–387.

Aronson, S. M. 1964 Epidemiology. In: Tay-Sachs' Disease. B. W. Volk, ed. Grune and Stratton, New York, pp. 118–154.

Buies, A. 1896 Le Saguenay et le Bassin du Lac Saint-Jean. Léger Brousseau, Québec, 3rd edition.

Dean, G. 1963 The prevalence of the porphyrias. So. Afr. J. Lab. Clin. Med., 9: 145–151.

Eriksson, A. W. 1961 Eine neue Blutersippe mit v. Willebrand-Jürgens' scher Krankheit (erbliche Thrompathie) auf Åland (Finnland). Acta Genet. Med. Gemell., 10: 157–180.

Ford, E. B. 1940 Polymorphism and taxonomy. In: The New Systematics. J. S. Huxley, ed. Clarendon Press, Oxford, pp. 493–513.

Gilbert, W. H. 1945 The wesorts of Maryland; an outcasted group. Jour. Wash. Acad. Sci., 35: 237–246.

Harte, T. J. 1959 Trends in mate selection in a tri-racial isolate. Social Forces, 37: 215–221.

———— 1963 Social origins of the Brandywine population. Phylon, 24: 369–378.

Henripin, J. 1954 La Population Canadienne au Début du XVIIIe Siècle. Institut national d'études démographiques, Paris, Cahier No. 22.

Herzog, M. I. 1965 The Yiddish Language in Northern Poland: Its Geography and History. Indiana University Research Center in Anthropology, Folklore and Linguistics, Pub. No. 37.

Knudson, A. G., L. Wayne and W. Y. Hallett 1967 On the selective advantage of cystic fibrosis heterozygotes. Amer. J. Hum. Genet., 19: 388–392.

Laberge, C., and L. Dallaire 1967 Genetic aspects of tyrosinemia in the Chicoutimi Region. Canad. Med. Assoc. J., 97: 1099–1100.

McKusick, V. 1966 Clinical genetics at a population level. The ethnicity of disease in the United States. Alabama J. Med. Sci., 3: 408–424.

McKusick, V., J. A. Hostetler, J. A. Egeland and R. Eldridge 1964 The distribution of certain genes in the Old Order Amish. Cold Spring Harbor Symp. Quant. Biol., 29: 99–114.

Morton, N. E. 1960 The mutational load due to detrimental genes in man. Amer. J. Hum. Genet., 12: 348–364.

Myrianthopoulos, N. C., and S. M. Aronson 1966 Population dynamics of Tay-Sachs Disease. I. Reproductive fitness and selection. Amer. J. Hum. Genet., 18: 313–327.

Roberts, D. F. 1965 Assumption and fact in anthropological genetics. Jour. Roy. Anthrop. Inst., 95: 87–103.

Rucknagel, D. L. 1964 The Gene for Sickle Cell Hemoglobin in the Wesorts. Thesis, University of Michigan, Ann Arbor.

Waldenström, J., and B. Haeger-Aronsen 1967 The porphyrias: a genetic problem. In: Progress in Medical Genetics. Volume V. A. G. Steinberg and A. G. Bearn, eds. Grune and Stratton, New York, pp. 58–101.

Witkop, C. J., C. J. MacLean, P. J. Schmidt and J. L. Henry 1966 Medical and dental findings in the Brandywine isolate. Alabama J. Med. Sci., 3: 382–403.