Am J Hum Genet 31:731–740, 1979

The Frequency of Private Electrophoretic Variants in Australian Aborigines and Indirect Estimates of Mutation Rate

K. K. BHATIA,[1] N. M. BLAKE, AND R. L. KIRK

SUMMARY

The number of "private" electrophoretic variants of enzymes controlled by 25 loci has been used to obtain estimates of mutation rate in Australian Aborigines. Three different methods yield values of 6.11×10^{-6}, 2.78×10^{-6}, and 12.86×10^{-6}/locus per generation for the total sample of Aborigines. One tribal population of Waljbiri in central Australia gives values of 2.99×10^{-6} and 2.04×10^{-6} for two of the methods, the third being unapplicable. The mean mutation rate for the total Aboriginal sample of 7.25×10^{-6} is very similar to the value obtained by Neel and his colleagues for Amerindians in South America.

INTRODUCTION

Several studies have used the frequency of private electrophoretic variants of blood proteins detected in samples from local human populations to indirectly estimate the average mutation rate per locus in man [1–5]. Neel and his colleagues and Tchen et al. based their calculations on data collected by themselves and collaborators from Amerindian populations in South America. Chakraborty and Roychoudhury relied on results published by workers from three laboratories, including our own, for tribal populations in India.

The formulations used and the basic data and assumptions needed in estimating the mutation rate from the frequency of rare variants have been detailed by Neel and Rothman [3]. They conclude that for electrophoretic variants the mutation rate averages 16×10^{-6}/locus per generation in Amerindian populations, but they point out that the possibility exists for variation in mutation rate on an ethnic or regional basis. Since this possibility requires exploration before statements can be made on the

Received March 26, 1979.

[1] All authors: Department of Human Biology, John Curtin School of Medical Research, Canberra, Australia.

© 1979 by the American Society of Human Genetics. 0002-9297/79/3106-0004$01.00

average mutation rate for man as a whole, we are analyzing our own extensive data for populations in the Southwest Pacific and Australian regions. We report here results for the Aboriginal populations of Australia. Results will be published later for populations in Papua, New Guinea, and for other parts of South and Southeast Asia and the Pacific.

THE STUDY POPULATION

At the time of first European contact, the Aborigines were spread across the Australian continent, having exploited, with few exceptions, all the available ecological situations. Their presence in the continent is dated back to at least 40,000 years, though the occupation of the more arid areas in the center probably took place no more than 10,000 years ago [6]. At the time of European contact the population of Aborigines has been estimated at about 250,000 [7], and the population was divided into several hundred tribal and local groups varying in size from 100 to several thousand persons [8].

During the last 200 years the Aboriginal population of Australia fell dramatically, reaching its lowest reported level in the census of 1921. This population decrease was not uniform; in some areas such as Tasmania, the eclipse was total, while in many others across the southern portion of the continent there are few, if any, persons of full Aboriginal descent remaining. In areas more remote from European settlement the decline in numbers was less, but even here the total may have been reduced to 50% before the increase in population characterizing the present situation commenced. The present analysis is based on samples from this area of minimum disturbance shown in figure 1.

There are no accurate records of the age structure in traditional Aboriginal populations. Available data refer to populations already exposed to varying degrees of European contact. At present, the age structure for persons of full Aboriginal descent shows a heavy-based pyramid with only 41.6% in the 15– 44 years age group [9]. In the traditional situation, each population may have varied in demographic parameters influenced by natural disasters such as prolonged drought or cyclones. Such factors may have led to drastic reductions in number followed by subsequent population expansion or by replacement through migration from neighboring groups. Over a longer time period, however, we assume that the population of the continent was in equilibrium, and that the average net increase was zero.

Since the precise boundary of the total Aboriginal population in our surveys is difficult to define, we have provided data also for one specific tribal group, defined by the spoken language Waljbiri, one of the largest linguistic groups in the Northern Territory. The Waljbiri territory (see fig. 1) covers 35,000– 40,000 square miles of arid and semi-arid country, and the population density averages one person per 25 – 27 square miles [10].

Meggitt's detailed study of the Waljbiri revealed that the dialectical Waljbiri tribe is further divided into four subgroups, namely, Yalpari (Lander), Waneiga, Walmalla, and Ngalia. Marriages between the subgroups are frequent, according to Meggitt. Tindale [11], however, found only 1.3% marriages between Ngalia and Walmalla, and no Yalpari-Waneiga marriages were recorded. Birdsell [12] claims that before 1935.

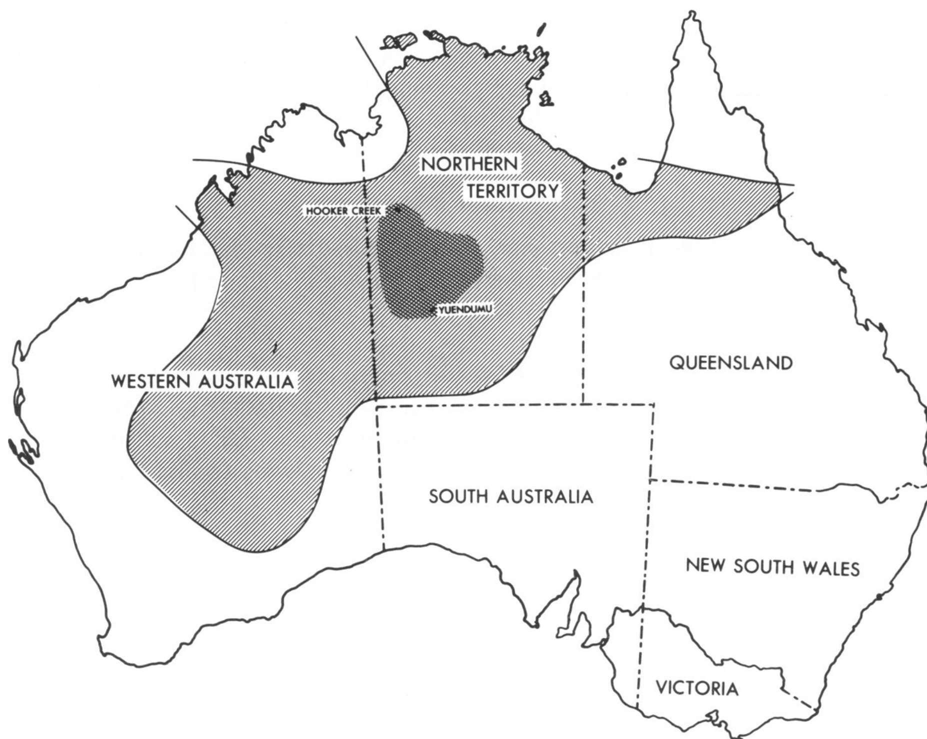

FIG. 1.—Map of Australia showing area sampled (*diagonal hatching*) and tribal territory of the Waljbiri (*cross hatching*).

the Ngalia subgroup was quite distinct from the other Waljbiri. Intertribal marriages involving Ngalia, however, were significantly higher at 6% – 7%.

The Waljbiri in our series were sampled mainly at two localities, Yuendumu and Hooker Creek. The Yuendumu Waljbiri predominantly belong to the Ngalia subgroup, though some reside also at Hooker Creek. Although we have pooled the results for all Waljbiri, our data indicate a clear-cut heterogeneity between the populations at these two localities.

THE LABORATORY DATA

Our analysis is confined to data for red cell enzyme proteins and hemoglobin, representing products of genes at 25 loci. The basic data have been tabulated recently by Blake [13] and are summarized in table 1. A total of 16 detected variants restricted to Australian Aborigines are listed in table 2, together with the number of copies observed and their gene frequencies. Three of the variant alleles ($PGM_2{}^3$, $CA_1{}^9$, and $CA_2{}^4$) have achieved frequencies above 1% and can be classified as polymorphic. Two others (PGD^{Elcho} and $PEP\ B^6$) have allele frequencies approaching 1%, and the remainder are more restricted, the number of copies ranging from one to 14. Table 2 also shows separately the number of rare variants detected in the Waljbiri tribe. Only five of the 16 rare variants among Aborigines were detected among the Waljbiri, four of

these being polymorphic in this tribe, while the other (*PEP* B^6) has an allele frequency of 0.74%. Three of the polymorphic alleles among the Waljbiri are polymorphic in Aborigines in general. In the case of the other peptidase variant allele (*PEP* B^7), 13 of the 14 copies occurred among Waljbiri, the other example being found in Luridja, a group known to intermarry with the Waljbiri.

METHODOLOGY

So far, three methods of indirectly calculating the mutation rate for electromorphs have been suggested and have been reviewed by Neel and Rothman [3].

Kimura and Ohta's Method

In Kimura and Ohta's method [14], three parameters are involved: the average number of mutant alleles per locus (I) estimated from all variants known to be restricted to the study populations; the effective size of the population (N_e), and the average mutant survival time in generations ($\bar{t}_o$) for alleles not moving toward fixation. The mutation rate, μ, is given by

$$\mu = \frac{I}{2N_e} \times \frac{1}{\bar{t}_o} .$$

TABLE 1

GENETIC MARKERS IN AUSTRALIAN ABORIGINES

Locus no.	Enzyme system	Abbreviation	Sample size
1	6-Phosphogluconate dehydrogenase	*6PGD*	4035
2	Acid phosphatase-1	ACP_1	4016
3	Phosphoglucomutase-1	PGM_1	3919
4	Phosphoglucomutase-2	PGM_2	3790
5	Peptidase A	*PEPA*	3034
6	Peptidase B	*PEPB*	3189
7	Carbonic anhydrase-1	CA_1	3751
8	Carbonic anhydrase-2	CA_2	3751
9	Glyoxylase	*GLO*	1290
10	Adenosine deaminase	*ADA*	1437
11	Esterase D	*EsD*	1556
12	Glutamic pyruvic transaminase	*GPT*	1391
13	Hemoglobin-α	$Hb\alpha$	2692
14	Hemoglobin-β	$Hb\beta$	2692
15	Diaphorase	*DIA*	1861
16	Glucose-6-phosphate dehydrogenase	*G6PD*	1014
17	Malate dehydrogenase-2	MDH_2	2964
18	Superoxide dismutase	*SOD*	1795
19	Lactate dehydrogenase-A	LDH_A	4180
20	Lactate dehydrogenase-B	LDH_B	4180
21	Isocitrate dehydrogenase	ICD_s	1226
22	Phosphohexose isomerase	*PHI*	1569
23	Adenylate kinase-1	AK_1	3535
24	Phosphoglycerate kinase	*PGK*	1569
25	Glutamic oxaloacetic acid transaminase	*GOT*	748

NOTE.—Based on Blake, 1979 [13].

TABLE 2

NUMBER AND FREQUENCIES OF PRIVATE VARIANTS IN AUSTRALIAN ABORIGINES

	Enzyme	Variant	Total Population		Waljbiri	
			No. copies	% Gene frequency	No. copies	% Gene frequency
1	6PGD	PGD^{Elcho}	65	0.81	0	0.00
2	PEPA	$PEP\ A^3$	1	0.02	0	0.00
3	PEPB	$PEP\ B^6$	53	0.83	6	0.74
4	PEPB	$PEP\ B^7$	14	0.21	13	1.43
5	PGM_1	PGM_1^6	4	0.05	0	0.00
6	PGM_1	PGM_1^7	1	0.01	0	0.00
7	PGM_2	PGM_2^3	103	1.36	46	5.68
8	PGM_2	PGM_2^{11}	6	0.08	0	0.00
9	ACP_1	ACP_1^F	1	0.01	0	0.00
10	CA_1	CA_1^9	192	2.53	36	4.44
11	CA_1	CA_1^{10}	8	0.11	0	0.00
12	CA_2	CA_2^4	166	2.21	14	1.73
13	LDH_B	LDH_B^{SLOW}	1	0.01	0	0.00
14	LDH_A	LDH_A^{SLOW}	2	0.02	0	0.00
15	G6PD	Gd_B^{FAST}	1	0.04	0	0.00
16	PHI	PHI^4	1	0.03	0	0.00

Nei's Method

Nei [2] gives a different formulation, where

$$\mu = \frac{I_q}{2N_e} \times \frac{1}{2 \log_e (2nq)} ,$$

and in this case, I_q is estimated from only those variants whose frequency does not exceed 1%, n is the sample size, and q is set at a value of 0.01.

Rothman and Adams' Method

Rothman and Adams [15] give

$$\mu = \frac{\hat{I}}{2N_e} \times [\bar{g}(1) - \Sigma\bar{g}(i)P_{i1}] ,$$

where $\hat{I}$ is the expected number of mutant alleles in the effective population per locus as estimated from the sample, $\bar{g}(1)$ is the estimated proportion of alleles present in only a single copy, $\bar{g}(i)$ is the estimated proportion of variant alleles of i copies, and P_{i1} is the probability for an allele represented by a single copy having i copies in the previous generation. The relationship between I and $\hat{I}$ is given by Rothman and Adams [15].

ESTIMATION OF I, I_q AND $\hat{I}$

For the total Aboriginal population, 25 enzyme loci have been studied. Of these, 13 showed no polymorphism by the standard definition where the least common allele frequency did not exceed 1%, but private variants were detected at four of these loci. Twelve private variants were distributed among eight of the 12 polymorphic loci. The mean sample size (table 3) for the polymorphic loci without private variants (1419) is

TABLE 3

NUMBER OF LOCI WITH AND WITHOUT PRIVATE VARIANTS AND VALUES OF I IN THE TOTAL ABORIGINAL POPULATION

	POLYMORPHIC LOCI			MONOMORPHIC LOCI			TOTAL LOCI		
	With private variants	Without private variants	Total	With private variants	Without private variants	Total	With private variants	Without private variants	Total
No. loci	8	4	12	4	9	13	12	13	25
No. private variants	12	0	12	4	0	4	16	0	16
I/locus	1.50	0.00	1.00	1.00	0.00	0.31	1.33	0.00	0.64
I_q/locus	1.12	0.00	0.75	1.00	0.00	0.31	1.08	0.00	0.52
$\hat{I}$/locus	2.07	0.00	1.38	4.93	0.00	1.52	1.80	0.00	1.07
Mean sample size/locus	3686	1419	2930	2736	2120	2309	3369	1904	2607
Mean sample size/variant	2457	...	2930	2736	...	7506	2527	...	4074

significantly lower than those with private variants (3686). In our data, therefore, the probability of detecting private variants among known polymorphic loci increases with sample size.

The data in table 3 clearly show that the value of I varies also with the type of loci (polymorphic or monomorphic). The probability of detecting private variants increases with sample size, which will also influence the value of I.

In Nei's method [2], I_q is calculated only from those private variants which are not polymorphic. The differences for all loci between I and I_q is 0.12 for the total Aboriginal population. For the Waljbiri population, the only private variant (*PEP B*^{7}) is polymorphic. Since it is an exclusive tribal marker for the Waljbiri, we have used it in the calculation of I_q, giving a value of 0.04.

Rothman and Adams' [15] method gives higher values for $\hat{I}$ than for either I or I_q, since the number of private variants is estimated for the total effective population. The difference is most marked for the monomorphic loci, where $\hat{I}$ is almost five times the value of I or I_q.

THE ESTIMATION OF N_e

As explained earlier, with the data available, it is not possible to give a precise estimate of the "effective" population size because of changing reproductive patterns among Aborigines. Here we use the population in the 15–44 age group in the 1961 census year, adjusted for the proportion of the total population in the surveyed area.

The population of full descent Aborigines in Australia in 1961 was 36,137 (18,899 males; 17,238 females), of which 41.6% were in the age cohort 15– 44 years [9], and the area surveyed contains approximately 60% of the total full descent population. This gives a value of $N_e = 9{,}160$.

It can be argued that this does not represent the effective population size of Australian Aborigines during most of their stay on the continent. However, indirect evidence suggests the difference in age structure in traditionally oriented societies is not likely to be very different from the value used here. For example, Tindale [8] has recorded approximate age composition for three nomadic bands encountered in the central desert areas. The mean value for the adult composition of these bands is 28.0%. Since this covers the age range 20– 40 years, the composition of the 15– 44 cohort will not be very different from the 41.6% derived from the 1961 census. In the case of the Waljbiri, we have age estimates for the Yuendumu population [16]. This gives 43.2% for the 15– 44 age cohort. From the total Waljbiri population estimate given by Milliken [17], N_e for Waljbiri becomes 1,173.

Another difficulty is that Aboriginal populations have been subject to a series of bottleneck effects due to the operation of various factors. This could cause the loss of a number of private variants which, in turn, will affect the calculation of I, I_q, and $\hat{I}$. The loss of these private variants, however, will be proportional to the decline in population size. On the other hand, private variants which survive the population crash will increase in number during the subsequent population expansion.

ESTIMATION OF $\bar{t}_0$

Kimura and Ohta [14] showed that the mean survival time for a neutral mutation in

generations in terms of effective population size and total population size is given by

$$\bar{t}_o = 2 \frac{N_e}{N} \log_e (2N) \ ,$$

in a stationary nonsubdivided population with no reproductive death and the progeny size following a Poisson distribution. Kimura and Maruyama [18], however, argue that if the population is subdivided into loose random mating units between which migration occurs, it may be treated approximately as a single random mating unit, disregarding the substructure of the populations.

Applying Kimura and Ohta's formula to the Aboriginal populations, and using the estimates of N_e given above, we obtain values of $\bar{t}_o$ = 10.7 and 7.0 generations for the total Aboriginal population, and Waljbiri, respectively. Neel and Rothman [3] however, consider the values of $\bar{t}_o$ calculated by this method as overestimates. The mean survival time can be simulated for each population, and Li and Neel [19] and Li et al. [20] obtained values between 2.3 to 2.8 generations. However, after making concessions for the various factors influencing the population, Li and Neel [19] believe a mean value of 5.7 generations is more appropriate. We shall use this value here.

ESTIMATION OF MUTATION RATES

Mutation rates estimated by each of the three methods listed above, both for the total Aboriginal population surveyed and for the Waljbiri tribal group are given in table 4. The rates vary within a range from 2.78×10^{-6} to 12.86×10^{-6}, with a mean value of 7.25×10^{-6}/locus per generation. In obtaining the values of μ based on Kimura and Ohta's [14] and Nei's [2] methods, we estimated the mean number of variants per locus using the sample size. Neel and Rothman [3], however, base their estimate on the total effective population size. Accordingly, we have also calculated μ with $I = \hat{I}$ and $I_q = \hat{I}$, and the new values become 10.18×10^{-6} and 5.72×10^{-6}/locus per generation, respectively.

Mutation rates estimated from private variants at polymorphic loci are 2.5 – 5 times higher than those estimated from the monomorphic loci. This may be a function of the smaller sample sizes for the private variants at monomorphic loci in our sample, which will have reduced the probability of detecting private variants. Eanes and Koehn [21] recently have also drawn attention to the relationship between sample size and detection of rare electrophoretic variants. Neel and Rothman's [3] method, however, yields a higher value of μ when all loci are considered together. The other two methods give values of μ for all loci intermediate between the values for polymorphic and nonpolymorphic loci, while in Neel and Rothman's method, the value of μ for all loci is higher than for either the polymorphic or nonpolymorphic loci.

The values of $\mu = 2.99 \times 10^{-6}$ and 2.04×10^{-6}/locus per generation for the Waljbiri, using Kimura and Ohta's and Nei's methods, are lower than the values obtained for the total Aboriginal sample. These lower values are due to the fact that while five private variants were detected in the Waljbiri, only one is included for calculating I and I_q. The other four are more widely distributed in the Aboriginal population, and it is not possible to assign the original mutants to the Waljbiri.

TABLE 4

MUTATION RATES ($\times 10^6$) IN AUSTRALIAN ABORIGINES ESTIMATED BY VARIOUS METHODS

	Kimura and Ohta's method I	Nei's method I_q	Rothman and Adams' method $\hat{I}$
Polymorphic loci, total Aboriginal sample	9.55	4.01	11.52
Monomorphic loci, total Aboriginal sample	2.96	1.66	2.28
All loci*, total Aboriginal sample	6.11†	2.78†	12.86
Waljbiri sample	2.99	2.04	...

* Mean = 7.25×10^{-6}.
† Following Neel and Rothman [3], the values of μ ($I = I_q = \hat{I}$) were estimated to be 10.18×10^{-6} and 5.72×10^{-6}, respectively.

Neel and Rothman [3] estimated mean mutation rates based on values for 12 Amerindian tribes in South America by each of the same three methods. The unweighted mean for the 12 tribes averaged for the three methods is 16×10^{-6}/locus per generation. The mean value for the Amerindians is more than twice our value for the total Aboriginal sample. However, Neel and Rothman's value of 16×10^{-6} is based on unweighted means for the tribal samples. If it is recalculated using weights based on the effective population sizes, the weighted mean value becomes 7.2×10^{-6}/locus per generation. It is interesting to note that recently Neel and Thompson [22], using a method based on simulation, arrived at a mean mutation rate of 7.0×10^{-6}/locus per generation. These values are very similar to our own based on the total Aboriginal sample. The value for the Waljbiri, of course, is only one-half that for the total Aboriginal sample. Neel and Rothman found a range of values of $0-51 \times 10^{-6}$/locus per generation for their 12 Amerindian tribes. The Waljbiri, therefore, fall within this range and we assume that if data were available for a similar number of tribal populations in Australia, the range of values may also be similar to those for the Amerindians.

Although the indirect estimation of mutation rates using data on private electrophoretic variants has many problems ranging from the technical factors influencing the recognition of rare variants through sampling design to the estimation of I, I_q, $\hat{I}$, and N_e, it is of great interest that data collected in two different laboratories from studies of different populations on two continents have yielded estimates of μ which are so similar. Further studies are in progress in our laboratory which we hope will make possible a further critical evaluation of this approach to the estimation of human mutation rates.

REFERENCES

1. NEEL JV: Private genetic variants and the frequency of mutation among South American Indians. *Proc Natl Acad Sci USA* 70:3311–3315, 1973
2. NEI M: Estimation of mutation rates from rare protein variants. *Am J Hum Genet* 29:225–232, 1977
3. NEEL JV, ROTHMAN ED: Indirect estimates of mutation rates in tribal Amerindians. *Proc Natl Acad Sci USA* 75:5585–5588, 1978

4. Chakraborty R, Roychoudhury AK: Mutation rates from rare variants of proteins in Indian tribes. *Hum Genet* 43:179– 183, 1978
5. Tchen P, Séger J, Bois E, Grenand F, Fribourg-Blanc A, Feingold N: A genetic study of two French Guiana Amerindian populations II: rare electrophoretic variants. *Hum Genet* 45:317– 326, 1978
6. Kirk RL: *Man Adapting: Human Biology of Australian Aborigines*. Oxford, Oxford University Press. In press, 1979
7. Radcliffe-Brown AR: Former numbers and distribution of the Australian Aborigines. *Official Yearbook of the Commonwealth of Australia* 23:671 – 696, 1930
8. Tindale NB: *Aboriginal Tribes of Australia*. Los Angeles, University of California Press, 1974
9. Commonwealth of Australia. *Population and Australia*. Canberra, Government Printing Service, 1975
10. Meggitt MJ: *Desert People*. Sydney, Australia, Angus and Robertson, 1962
11. Tindale NB: Tribal and intertribal marriage among the Australian Aborigines. *Hum Biol* 25:169– 190, 1953
12. Birdsell JB: Local group composition among the Australian Aborigines: a critique of the evidence from fieldwork conducted since 1930. *Curr Anthropol* 11:115– 142, 1970
13. Blake NM: Genetic variation of red cell enzyme systems in Australian Aboriginal populations. *Occasional Papers in Human Biology* 2:39– 82, Canberra, Australian Institute of Aboriginal Studies, 1979
14. Kimura M, Ohta T: The average number of generations until fixation of a mutant allele. *Genetics* 61:763– 771, 1969
15. Rothman ED, Adams J: Estimation of expected number of rare alleles of a locus and calculation of mutation rates. *Proc Natl Acad Sci USA* 75:5094– 5098, 1978
16. Middleton MR, Francis SH: *Yuendumu and its children*. Canberra, Australian Government Publishing Service, 1976
17. Milliken EP: Aboriginal language distribution in Northern Territory, in *Tribes and Boundaries in Australia*, edited by Peterson N, Canberra, Australian Institute of Aboriginal Studies, 1976
18. Kimura M, Maruyama T: Pattern of neutral polymorphism in a geographically structured population. *Genet Res* 18:125– 131, 1971
19. Li FHF, Neel JV: A simulation of the fate of a mutant gene of neutral selective value in a primitive population, in *Computer Simulation in Human Population Studies*, edited by Dyke B, MacCluer JW, New York, Academic Press, 1974, 221 – 240
20. Li FHF, Neel JV, Rothman ED: A second study of the survival of a neutral mutant in a simulated Amerindian population. *Am Naturalist* 112:83– 96, 1978
21. Eanes WF, Koehn RK: Relationship between subunit size and number of rare electrophoretic alleles in human enzymes. *Biochem Genet* 16:971 – 985, 1978
22. Neel JV, Thompson EA: Founder effect and number of private polymorphisms observed in Amerindian tribes. *Proc Natl Acad Sci USA* 75:1904– 1908, 1978