Albumin Yanomama-2, a 'private' polymorphism of serum albumin

BY ROBERT TANIS, ROBERT E. FERRELL, JAMES V. NEEL AND
MARIANNE MORROW

*Department of Human Genetics, University of Michigan Medical School,
Ann Arbor, Michigan* 48104

INTRODUCTION

At least 26 variants of human serum albumin have been demonstrated by suitable electrophoretic techniques (reviews in McDermid, 1971; Weitkamp *et al.* 1973*a*, *b*). In addition to the rare albumin variants, by now encountered in most of the major ethnic groups, at least two instances of an albumin variant occurring in polymorphic proportions have been reported. The first such variant, Albumin Naskapi, was initially encountered by Melartin & Blumberg (1966) as a rapidly migrating albumin fraction transmitted as if due to a co-dominant gene which had a frequency of 0·13 in a sample of 151 Naskapi Indians of Quebec Province, Canada, and lower frequencies within other North American Indian populations (the Montagnais, Sious, Tlingit, and Athabascan speakers from Alaska). Later this variant was encountered in three tribes of the Southwestern United States (the Navajo, Apache and Hopi) (Johnston *et al.* 1969); however, it has not been observed in Central or South American Indians. Another well-documented serum albumin variant also occurring as a polymorphism has been designated Albumin Mexico by Melartin and co-workers (1967). This is a slowly migrating variant originally detected with a gene frequency of 0·03 in a sample of 185 'mestizos' from Tlaxcala and Hidalgo states and Mexico, D.F. Later it was found in the Papago, Navajo, Apache, and Pima Indians of the Southwestern United States and the Quiche of Guatemala (Johnston *et al.* 1969, 1973); the gene frequency for albumin Mexico amongst these tribes ranged from 0·006 to 0·049.

A recent report by Tanis *et al.* (1973) concerned itself in part with describing eight serum albumin variants detected in five tribes of South American Indians residing in the north-central portion of that continent. In all cases but one, the electrophoretic variants were rare. However, the variant designated Yanomama-2 (Yan-2) was detected in 46 individuals within a rather limited geographic area occupied by the Yanomama Indians, a tribe of more than 12,000 persons residing in Northern Brazil and Southern Venezuela. Weitkamp and colleagues (1973*a*) have compared Albumin Yan-2 with some 23 other serum albumin variants under a variety of starch-gel conditions; under the conditions employed its migration behaviour appears unique.

The Yan-2 variant did not separate well from normal in the test systems in use in this laboratory at the time of its discovery (see Materials and Methods). In an effort to improve upon this situation, a new albumin screening system has been developed. Because of the possibility that many examples of the Yan-2 variant could have been missed during the earlier work from this laboratory, due to lack of adequate resolution by the two albumin screening systems employed routinely, we have retyped all the samples of Indian serum available in this Department, some 5993 in number from a total of 11 tribes. We now find that the Yan-2 variant is widespread throughout the Yanomama, exhibits a geographic cline, and is not found in the members of any other tribe thus far studied.

The techniques of collection and storage of the blood samples from the Indians involved in this report have been treated adequately elsewhere (Gershowitz *et al.* 1972). Samples from the Yanomama, Makiritare, Piaroa, Macushi, Wapishana and Xavante Indian tribes were collected by members of the Department of Human Genetics of the University of Michigan during the period 1964–72. Samples from the other five tribes included in this study were supplied by colleagues at other institutions, as indicated in Table 3, to all of whom we are most grateful. A reference sample of Albumin Mexico was provided by Dr H. E. Sutton and of Albumin Naskapi by Dr. D. L. Rucknagel.

The previous screening of these samples for albumin variants relied primarily on starch-gel electrophoresis systems employing a pH 5 sodium acetate buffer or a pH 6·9 Tris–EDTA–boric acid buffer as described by Weitkamp and co-workers (Weitkamp *et al.* 1969, 1973a). With these systems, the albumin of a heterozygote for the Yan-2 variant is typically seen as a relatively broad band. To obtain the resolution necessary for the reliable detection of both the heterozygote and the homozygote patterns for this gene, we have developed an improved starch-gel system. Thus far the clearest resolution is obtained with vertical starch-gel electrophoresis (Electro Starch; Otto Hiller, Madison, Wisconsin), employing a pH 6·1 gel buffer containing 4×10^{-3} M Tris, 5×10^{-4} M EDTA and 6×10^{-2} M boric acid; the bridge buffer, also pH 6·1, consists of undiluted stock (3×10^{-2} M Tris, 4×10^{-3} M EDTA and 0·5 M boric acid) on the gel side, while one volume each of stock buffer, 20 % NaCl, and H_2O were used on the electrode side. The reader is referred to the treatment by Brewer (1970) for a detailed explanation of the equipment used. Optimum electrophoretic conditions were found to be 7 V./cm. for 6 hr. at 4° C. Using the above conditions both the heterozygous and homozygous patterns of albumin Yan-2 could easily be distinguished from each other and from the normal albumin pattern (Fig. 1, wells 1–3). However, to confirm that all 30 of the samples thought to exhibit only Yan-2 albumin were validly typed, each was mixed with an equal volume of normal serum and the mixture subjected to electrophoresis; in each case the 'bis' pattern was identical to the observed pattern of the heterozygotes (compare well 3 with well 4 (mixture), Fig. 1).

Although the new buffer system was developed specifically to improve resolution for the Yan-2 variant, it was observed to result in an adequate separation from normal of several other of the albumin variants of American Indians which have been previously reported. During the course of rescreening the 5993 samples, the following previously observed albumin variants were again detected: Mak-3, Wapishana, Maku, Mexico and Naskapi. The electrophoretic behaviour of each, using the pH 6·1 TEB buffer, is also shown in Fig. 1. It should be noted that albumins Maku and Wapishana are presently considered identical electrophoretically and therefore for the purposes of this paper are treated as a single variant designated Maku. For the present, we are forced to equate electrophoretic similarity of two albumins with biochemical identity but the experience with the abnormal haemoglobins makes it clear that with technical advances, we may anticipate establishing differences between some molecules with identical mobilities. The reader is referred to the work of Tanis *et al.* (1973) and Weitkamp *et al.* (1973a) for a more detailed treatment of many of these albumin variants. We further noted that this new buffer system, although providing excellent resolution of monomeric variants of albumin, is unsuitable for the detection of dimeric variants.

(+)

1 2 3 4 5 6 7 8 9 10 11 12 ←

Fig. 1. Starch-gel patterns of the several albumin variants seen during the course of this study. From left to right: Well 1, Normal; Well 2, Homozygous Albumin Yanomama-2; Well 3, Heterozygous Albumin Yanomama-2; Well 4, mixture of normal and homozygous Albumin Yanomama-2; Well 5, Albumin Mexico; Well 6, Albumin Naskapi; Well 7, normal; Well 8, Albumin Makiritare-2; Well 9, Albumin Makiritare-3; Well 10, Heterozygous Albumin Yanomama-2; Well 11, Albumin Wapishana; Well 12, Albumin Maku. The sample labelled Makiritare-2 was not ascertained through our customary sampling procedures and hence is not included in Table 3; it migrates identically with Maku and Wapishana. The origin is indicated by an arrow.

RESULTS

The phenotype and gene frequencies for albumin Yan-2 in the Yanomama Indians are presented in Table 1; the location of most of the villages is shown in Fig. 2. These are the same villages for which we have presented gene frequencies for 28 other systems in previous publications from this group. We note that the frequency of the Yan-2 gene in individual villages varies from 0·0 to 0·40.

Table 1. *Phenotypic totals and gene frequencies for Albumin Yanomama-2 in 47 major villages† and 17 miscellaneous villages of the Yanomama Indians*

Subdivision	Village	N	$N + Y_2$	Y_2	Σ	$F(Y_2)$	Other albumin variants
I	03X	68	3	0	75	0·020	4‡
	03W	60	5	0	65	0·039	
	03T	51	2	0	53	0·019	
	15L	67	14	0	81	0·086	
	Misc.						
	03Y	13	3	0	16		
	03Z	7	0	0	7	0·066	
	15O	22	2	0	24		
	15P	4	2	0	6		
	Total	292	31	0	327	0·047	
II	03LMN	66	4	0	70	0·028	
	03KP	61	13	0	74	0·088	
	03Q	33	2	0	35	0·029	
	03RS	68	1	0	69	0·007	
	Total	228	20	0	248	0·040	
III	03U	18	24	2	44	0·318	
	08D	17	11	3	31	0·274	
	08E	11	24	3	38	0·395	
	08F	12	16	4	32	0·375	
	08I	54	8	0	62	0·064	
	08XY§	114	7	0	121	0·029	

R. TANIS AND OTHERS

Table 1 (cont.)

Subdivision	Village	N	$N + Y_2$	Y_2	Σ	$F(Y2)$	Other albumin variants
III	11ABC	149	1	0	150	0·003	
	11D	27	20	3	50	0·260	
	11S	26	12	1	39	0·179	
	11X	40	5	0	45	0·056	
	15H	39	4	0	43	0·046	
	Misc.						
	03V	19	2	0	21		
	08G	6	1	0	7		
	08H	2	4	1	7		
	08Z	20	1	0	21	0·088	
	11V	18	10	1	29		
	11W	11	0	0	11		
	15G	19	0	0	19		
	15I	24	3	0	27		
	Total	626	153	18	797	0·120	
IV	03J	57	4	0	61	0·033	
	08J	30	1	1	32	0·047	
	08K/15M	75	6	0	81	0·037	
	08L/11P	87	4	0	91	0·022	
	08O	29	6	0	35	0·086	
	08Q	43	5	0	48	0·052	
	08R	51	6	0	57	0·053	
	11M	31	0	0	31	0·0	
	Misc.						
	08P	4	6	0	10		
	11N	9	5	0	14	0·221	
	'15N'	27	15	2	44		
	Total	443	58	3	504	0·064	
V	03A	61	19	1	81	0·130	
	03B	39	4	0	43	0·046	
	03E	49	11	0	60	0·092	
	03F	34	0	0	34	0·0	
	03G	34	14	0	48	0·146	
	03I	45	2	3	50	0·080	
	08ABC	147	26	3	176	0·091	
	08N	41	9	0	50	0·090	
	08S	55	9	0	64	0·070	
	08T	24	6	1	31	0·129	
	08UVW	143	19	0	162	0·059	
	11K	38	8	0	46	0·087	
	11T	26	6	0	32	0·094	
	Misc.						
	08M	3	1	0	4	0·078	
	11U	24	4	0	28		
	Total	763	138	8	909	0·086	
VI	03C	50	9	0	59	0·076	
	03D	76	0	0	76	0·0	
	03H/11J	104	2	0	106	0·010	
	11G	38	0	0	38	0·0	
	11HI	123	5	0	128	0·020	
	11YZ	112	18	0	130	0·069	
	15QR	88	27	1	116	0·125	
	Total	591	61	1	653	0·048	
	Grand total	2943	461	30	3438	0·076	4

† Explanation in text.

‡ Albumin Maku (Weitkamp & Chagnon, 1968).

§ The 'village' designated 08XY is a cluster of three closely related villages, shown as 15D, E and F in Fig. 3.
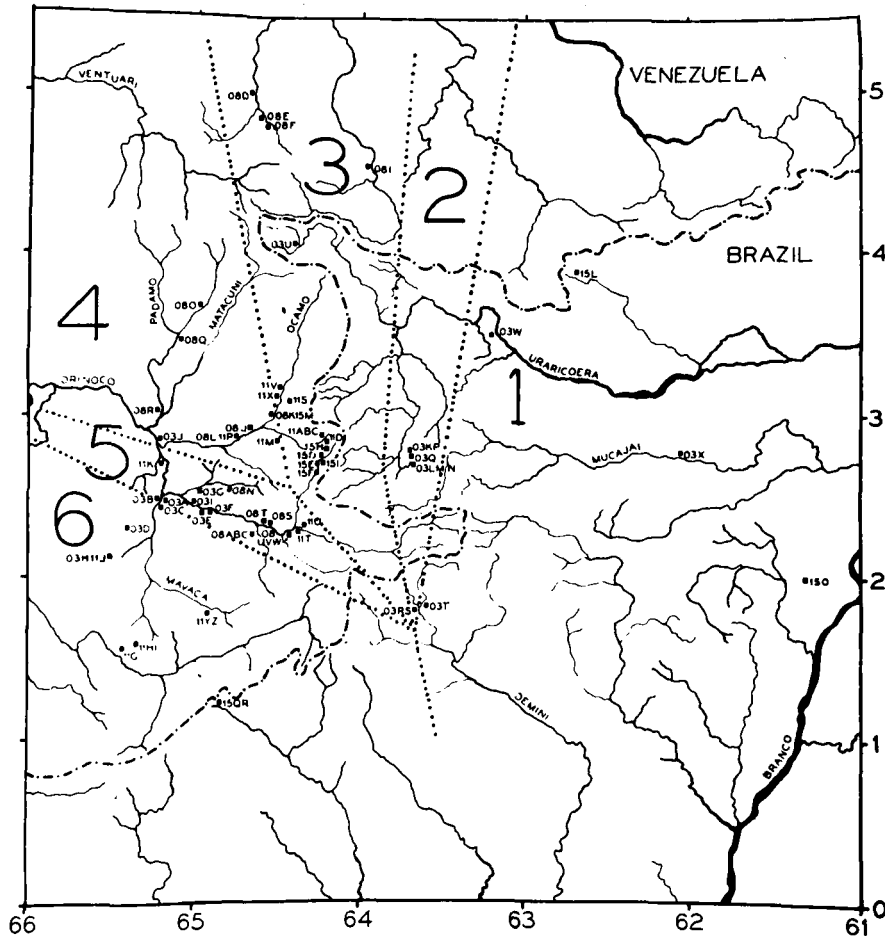
Fig. 2. Map showing the geographical location of 53 Yanomama villages from which biological specimens have been collected. Superimposed on the map are the six subdivisions defined by genetic cluster techniques (Neel *et al.* 1972). Villages 15D, 15E and 15F are represented in Table 1 as the composite village 08XY.

A question of interest is whether there are significant local differences in the frequency of the Yan-2 gene within the distribution of the Yanomama. In previous publications, we have shown that on the basis of classification by 11 genetic systems (largely blood types and serum proteins, and *not* including albumin), cluster analysis divided 27 Yanomama villages into six reasonably defined groups (Ward, 1972; Neel *et al.* 1972). These six groups occupied geographically non-overlapping areas, as shown in Fig. 2. The ten additional villages studied since this cluster analysis can all be assigned to one of these six areas. Table 1 presents the findings by village and geographic subdivision; the position of all 47 villages is shown in Fig. 2. One objective way to approach the question of local differences in gene frequency is to employ these six subdivisions as the basis for a comparison. Amongst subdivisions, the gene frequency varies from 0·040 to 0·120; a heterogeneity χ^2 for the gene frequency in the six subdivisions is significant ($\chi^2 = 79.37$, D.F. $= 5, P < 0.001$). (Because of the non-binomial nature of the data, created by the presence of biologically related persons, all significance tests must be regarded as approximate.) Thus, a genetic subdivision of the tribe by one set of criteria is reflected in the distribution of this additional genetic marker.
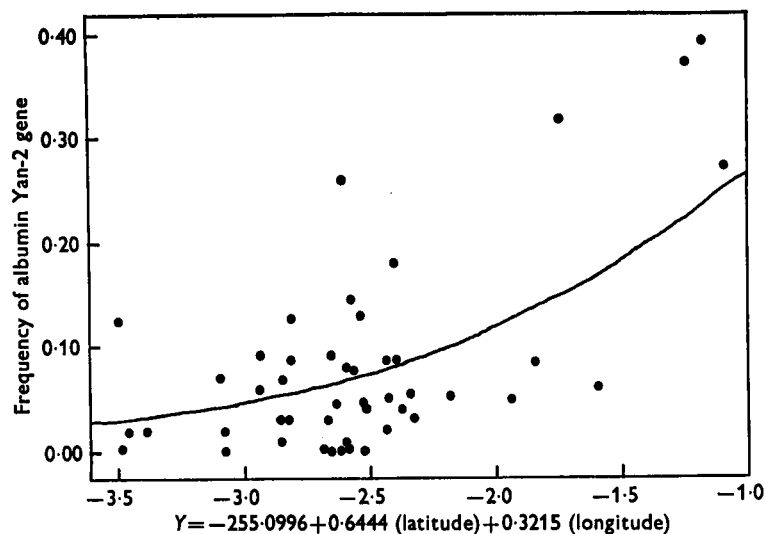
Fig. 3. Plot of village gene frequency of the Yanomama-2 allele in relation to latitude. The curved line represents the predicted gene frequency obtained by multinomial regression analysis.

Subdivision 3, with the highest mean gene frequency, is exceptionally heterogeneous in the gene frequencies of its constituent villages, with, for example, adjacent villages 11ABC and 11D differing by a gene frequency of 0·26. Neverthless, there seems to be a clear pattern within this sector. If we divide the area into two segments, above and below lat. 3° N, the gene frequency in the northern villages (03U, 03V, 08E, 08D, 08F, 08G, 08H, 08I and 11X) is 0·21, whereas in the southern villages (08XY, 08Z, 11ABC, 11D, 11S, 11V, 11W, 15G, 15H and 15I) it is 0·07. The uniqueness of sector 3 is therefore the result of the villages in the north since the frequency data for the southern villages are similar to that for the rest of the tribe.

A second approach to the question of local differentiation is to seek evidence of a cline. Because of the substantial area occupied by the tribe, geographic coordinates would appear to be convenient references in the search for clinal patterns. Using the regression procedures described by Smouse (1974), we obtained the estimated coefficients of the logistic regression equation:

$$Y = \log_e\left[\frac{P}{1-P}\right] = \beta_0 + \beta_1 \text{ (latitude)} + \beta_2 \text{ (longitude)}.$$

The observed values of the village gene frequencies are plotted, in Fig. 3, against the estimated regression equation. The abscissa of Fig. 3 can be thought of as a line running through the territory of the Yanomama from SE to NW at an angle of about 63·5° from the equator; the entries corresponding to the values of Y were yielded by the equation.

The statistical significance of the cline is examined in Table 2. The total variation among villages is partitioned according to the procedures of Smouse & Kojima (1972) into a regression (clinal) component and a residual (lack of fit) component. The regression component can be further subdivided as shown, and indicates that latitude is more useful as an indicator than longitude. The Λ criteria asymptotically approach χ^2 measures, with the indicated degrees of freedom. All criteria are highly significant, but since the various individuals within a village are often related (non-binomial sampling), one should view the probability levels with considerable caution. Inasmuch as the 'lack of fit' term (a measure of the inadequacy of the model) is quite large, one must also view the adequacy of the regression model with some reservation.

Table 2. *Analysis of gene-frequency variation among villages for the albumin locus*

Source of variation	D.F.	Λ†	$[\Lambda \div \text{D.F.}]$	F‡
Regression (cline)	2	118·82**	59·41	8·04**
Latitude	1	100·22**	100·22	13·57**
Longitude/latitude	1	18·60**	18·60	2·52[n.s.]
Lack of fit (resid.)	44	325·01**	7·39	
Total among population	46	443·83**		

** $P < 0.01$.   † Asymptotic χ^2.   ‡ Asymptotic F.

In this situation, a conservative procedure would be to gauge the regression component relative to the lack of fit. If we divide each Λ estimate by its degrees of freedom, we may compare the regression and lack of fit components on a per-degree-of-freedom basis. This is analogous to constructing mean squares from sums of squares in a conventional regression analysis of normal variates. Finally, since the Λ terms are asymptotically χ^2, and since our sample size is very large ($2n = 6278$), we may construct an approximate F-ratio from these two 'mean squares'. This value is 8·04, has $P < 0.01$, and even with allowance for non-independence of samples from within villages indicates a real trend.

The partitioned components of the regression χ^2 can be similarly treated; if this is done, we find that the latitudinal component is significant but that the longitudinal component is not. It is really the combination we are concerned with, and as indicated above, the combination of latitude and longitude is quite effective in describing the trend in gene frequency of the Yanomama-2 allele across the territory of the Yanomama.

Table 3 summarizes the results of screening 2562 serum samples from members of ten additional tribes from both South and North America with the new screening system. The variant labelled Maku encountered in 15 Wapishana is electrophoretically identical with the variant encountered in a captured Maku woman and three of her descendants living with the Yanomama (see Table 1). The Yan-2 variant was not encountered in any of these other samples. (In a sample of 66 blood specimens drawn from the Makiritare of Santa Maria de Erebato some time after the sample described in this paper was obtained, for the purpose of typing for the histocompatibility antigens, there was one specimen heterozygous for Yan-2. The representation of this sample cannot be related to our previous sample from this location, i.e. the amount of overlap is unclear, and we present only the finding of the first sample.) This absence of Yan-2 from other tribes is especially noteworthy in the case of the Makiritare; for at least the past 100 years this tribe has overlapped in its distribution with the Yanomama. This overlap is largely in the northern portion of the Yanomama territory where the frequency of the Yan-2 gene is especially high. The present study thus provides no evidence of gene flow from the Yanomama to these other tribes.

The reader is reminded that dimeric variants are not detected by the test system which generated the results of Tables 1 and 3. For the sake of the discussion to follow, we note the previous detection of two electrophoretically different dimeric variants in this material, namely Yanomama-1 and Makiritare-1, the latter electrophoretically indistinguishable from a variant we encountered in the Piaroa (Tanis *et al.* 1973), and also encountered by others in the Warao, Trio, and Wajana (Weitkamp *et al.* 1973*a*; Geerdink *et al.* 1974).

Table 3. *Distribution of albumin variants among ten American Indian tribes, excluding the Yanomama, as detected with the pH 6·1 starch-gel system*

(The Yanomama-1 and Makiritare-1 albumin dimers previously seen with other gel techniques (Tanis *et al.* 1973) are not demonstrated in this system.)

Tribe	Normal	Maku[a]	Makiritare-3[b]	Naskapi[c]	Mexico[d]	Total
Makiritare	717	0	2	0	0	719
Piaroa	146	0	0	0	0	146
Macushi	187	0	0	0	0	187
Wapishana	47	15	0	0	0	62
Xavante[e]	283	0	0	0	0	283
Cayapo[f]	580	0	0	0	0	580
Seminole-Creek[g]	186	0	0	0	0	186
Chippewayan[h]	85	0	0	2	0	87
Ojibwa[i]	112	0	0	8	0	120
Navaho[j]	177	0	0	11	4	192
Grand total	2520	15	2	21	4	2562

[a] Weitkamp & Chagnon (1968).
[b] Tanis *et al.* (1973).
[c] Melartin & Blumberg (1966).
[d] Melartin *et al.* (1967).
[e] The samples demonstrating the Xavante variant reported by Weitkamp *et al.* (1973 b) had been exhausted at the time of these repeat determinations.
[f] From Dr F. M. Salzano, Federal University of Rio Grande do Sul, Porto Alegre, Brazil.
[g] From Dr W. S. Pollitzer, University of North Caroline, Chapel Hill, North Carolina.
[h] From Dr Elizabeth Ives, University of Saskatchewan, Saskatoon, Canada.
[i] From Dr T. E. Reed, University of Toronto, Toronto, Canada.
[j] From Dr J. D. Niswander, National Institute of Dental Research, Washington, D.C.

DISCUSSION

The present study has defined an apparently unique situation, a gene which has achieved a mean frequency of 0·08 within a single Amerindian tribe without as yet evidence it has spread to adjacent tribes. This gene manifests itself as an electrophoretic variant of a protein molecule, serum albumin, for which seven other electrophoretically distinguishable variants have already been recognized in our samples from this major ethnic grouping, namely, the Yan-1 dimer, Mak-1 dimer, Mak-3, Maku, Naskapi, Mexico and Xavante. (Because of the way it was ascertained (cf. above and Tanis *et al.* 1973), we do not include the Makiritare-2 variant in this enumeration.) The Xavante variant, not previously mentioned, was reported to be unique by Weitkamp *et al.* (1973b), but does not appear in our Table 3 because the samples demonstrating this variant had been exhausted at the time of these repeat determinations. The range in the frequencies of the genes responsible for these various albumin variants is noteworthy. At the one extreme, there are variants apparently restricted to one or a few individuals within a single village of a single tribe, e.g. Yanomama-1, Xavante, and Makiritare-3 (Weitkamp *et al.* 1973b; Tanis *et al.* 1973). At the next level should be variants which occur with a low frequency in several villages of a single tribe; this situation has not yet been observed. Then there are variants which occur in polymorphic proportions within a single tribe, with no spill-over into adjacent tribes, such as the variant which is the subject of this paper. Finally, then, are variants in polymorphic proportions in one or several tribes, with significant but lesser representation in other tribes. This would appear to be true of albumins Naskapi and Mexico, and may also be true of

albumin Maku, although the data concerning the latter are still exiguous. Finally, we mention the Makiritare-1 dimer, found by us in 13 persons in four Makiritare villages in proportions bordering on a polymorphism, and in two Piaroa. This variant was also encountered by Arends *et al.* (1969) in 5 among 222 Warao of Venezuela and by Geerdink *et al.* (1974) in polymorphic proportions in the Trio and Wajana of Surinam. Again we indicate that as long as electrophoretic mobility remains the sole criterion for distinguishing between variants, any given 'variant' may well encompass several structural changes all characterized by the same apparent electrophoretic behaviour. The possibility of heterogeneity seems especially strong for dimeric variants such as Makiritare-1.

Elsewhere we have argued that the Yanomama as a tribe may have enjoyed an exceptional measure of isolation from all the surrounding tribes (Neel *et al.* 1972; Spielman, Migliazza & Neel, 1974; Ward *et al.* 1974). The present data are felt to reinforce this argument. The most common means whereby genes cross tribal boundaries is probably in consequence of hit-and-run attacks in which a raiding party, encountering a small group of traditional enemies in the field, kills the men and abducts the women and children. The present data suggest that few Yanomama women have contributed to the gene pool of the surrounding tribes.

Our previous studies have revealed microdifferentiation among Yanomama villages of the order of magnitude of that seen for Yan-2 with respect to a variety of other genetic systems (MNS, Rh, P, Duffy, Kidd, Lewis, Haptoglobin, GC, PGM_1 and GM), and, as we shall report elsewhere, five of those systems also exhibit significant clines. A standard problem for any possible explanation of this and the other clines will be the irregularities which they present. Accordingly, it seems appropriate at this juncture to recall the mobility of Yanomama villages. For instance, we have documented in detail a situation in which a Yanomama village migrated approximately 200 miles in the course of three generations (Chagnon *et al.* 1970). Given this mobility, and the fact that some Yanomama villages are essentially composed of 5–6 extended lineages, with the attendant opportunities for a founder effect, we feel that even in a situation where a cline is due to strongly deterministic pressures, irregularities are to be expected because of stochastic factors.

There are two obvious explanations of the frequency of the Yan-2 allele. One is the appearance through mutation at an early stage in tribal history of a gene with neutral (or even slightly disadvantageous) phenotypic effects, which has increased in frequency by chance alone (i.e. drift). The chief alternative explanation involves the appearance early in the history of the tribe of a gene with favourable effects, which has then increased to its present frequency by virtue of its selective advantage. Some combination of those two hypotheses is of course tenable. The occurrence of a cline is consistent with either of these explanations but a third explanation sometimes invoked to account for a cline, namely admixture, seems excluded in view of the evidence for the relative isolation of the tribe (Neel *et al.* 1972; Spielman *et al.* 1974).

We will on another occasion examine in detail the various kinds of evidence bearing on selection and the Yan-2 allele. Here, having summarized the present status of our knowledge of albumin variants in the American Indian, we wish only to point out that in reaching a decision concerning the dynamics of the albumin locus, the comparative approach may be of considerable value. This structural locus was chosen for a survey of the occurrence and frequency of alleles not because genetic polymorphisms were to be found at this locus, but simply because the protein product was suitable for our purposes. Eight electrophoretically distinguishable variants have

been observed in 5993 samples from members of eleven Indian tribes; 9·2% of the Indians sampled had an albumin variant (since the two samples demonstrating the Xavante-2 variant were exhausted by the time of this re-examination, the actual number of specimens examined was 5995). Since many of the Indians sampled are related to one another, the number of independent genomes sampled is of course substantially less than the heat count. By contrast, we have encountered no haemoglobin variants in 3595 samples from six of these same tribes* (Tanis *et al.* 1973). The work by others cited in the Introduction tends to confirm our finding of a relatively high frequency of albumin variants in the Amerindian. On the other hand, the summary of the results of the electrophoresis of haemoglobin samples from Amerindians by Arends (1971) reveals that aside from the S and C variants (probably introduced by admixture), only two certain and one possible, electrophoretically distinguishable haemoglobin variants (in a total of six persons) were encountered in 21,744 determinations. Thus, the lack of haemoglobin variants observed in our studies is by no means exceptional.

It appears that this difference between the albumin and haemoglobin loci may extend to another mongoloid group. As is well known, in recent years Japanese investigators have mounted an extensive search for variants of haemoglobin A. Iuchi (1968) reported that between 1957 and 1967 a total of 279,600 Japanese were screened, with the discovery of 44 variants (1:6400). The number of different electrophoretic behaviours is not given; structural analysis apparently reveals 35 different variants. The locations at which the screening was carried out are not given. At about the same time, Yanase and colleagues (1968) reported on a study of 50,000 haemoglobin specimens from successive clinic patients from the island of Kyushu; it is not clear whether this series is included in Iuchi's summary. Eighteen variants were encountered, which upon sequencing reduced to nine different mutants. There has thus far been no extensive search for albumin variants in Japan by methods comparable to our own. However, we (Ueda *et al.* 1974) have encountered a slow albumin variant in 5 of 2455 persons screened. The data thus suggest the same difference between albumin and haemoglobin in Japanese as in Amerindians.

Since the number of amino acids comprising albumin is approximately twice the number in the α and β chains of haemoglobin combined, one might, on the assumptions of equimutability per nucleotide and selective neutrality of the variants, anticipate twice as high a variant rate for the albumin locus as the α and β loci combined. Even with this 'correction', however, it is difficult to escape the impression that the frequency of occurrence of variants and the numerical representation of each at the three structural gene loci concerned are not comparable in these ethnic groups. The findings permit of two tentative interpretations: (1) if haemoglobin is taken as the reference molecule, the mutation rate at the albumin locus is higher and/or a larger fraction of the albumin variants is neutral or subject to positive selection, or (2) if albumin is taken as the reference molecule, the mutation rates at the α and β haemoglobin loci combined are lower and/or a larger fraction of the haemoglobin variants is subject to negative selection.

It is an obvious question, whether the same difference exists in Caucasians. The most extensive survey for haemoglobin variants in Caucasoids is that of Sick *et al.* (1967), who found 10 variants in 8000 persons, of which it appears that at least 5 and possibly 8 were electrophoretically distinguishable. Weitkamp *et al.* (1973a) reviewed the published data on albumin variants in

---

* In a related screening effort (undertaken by Dr P. Kale, private communication) a total of 2412 of our South American Indian haemolysates were re-examined for possible electrophoretic haemoglobin variants using a cellulose acetate methodology; no haemoglobin variants were observed.

Europeans, finding that 18 had been encountered in 43,382 determinations, but believed that for technical reasons this might be a substantial underestimate; it is not possible to determine how many different variants were represented among these 18. About all that can be said is that the data to date do not suggest the same striking difference between the two proteins that was encountered in the Indians and Japanese.

Significant progress is being made on the theory of the distribution of rare variants with neutral phenotypic effects in finite, subdivided populations (Ewens, 1972). Our present very limited data on Indian villages and tribes do not yet warrant a formal treatment of this subject, but work now in progress should result in a sufficient extension of the data so that we may in the near future be able to use this approach in an effort to determine whether the distribution of rare variants at the albumin locus is more consistent with that expected for mutations with neutral effects than is that of haemoglobin variants.

SUMMARY

A total of 5993 samples of serum or plasma from members of 11 Amerindian tribes have now been screened for the Yanomama-2 albumin variant, using improved techniques. Our results to date demonstrate that this albumin variant occurs only in a single South American Indian tribe, the Yanomama, with a gene frequency of 0·08; the frequency in individual villages ranges from 0 to 0·40. Significant geographic heterogeneity in the distribution of the gene is demonstrated by a comparison of its frequency in six previously established subdivisions of the tribe and also by the demonstration of a significant cline. The number and frequency of variants of albumin in the Amerindian is contrasted with the data for haemoglobin A; it is suggested that the interplay of mutation and selection differs at these two loci.

REFERENCES

ARENDS, T. (1971). Hemoglobinopathies and enzyme deficiencies in Latin American populations. In *The Ongoing Evolution of Latin American Population*, (ed. F. M. Salzano). Springfield, Ill.: Charles Thomas.

ARENDS, T., GALLANGO, M. L., LAYRISSE, M., WILBERT, J. & HEINEN, H. D. (1969). Albumin Warao: new type of human alloalbuminemia. *Blood* **33**, 414.

BREWER, G. (1970). *An Introduction to Isozyme Techniques*. New York: Academic Press.

CHAGNON, N. A., NEEL, J. V., WEITKAMP, L. R., GERSHOWITZ, H. & AYRES, M. (1970). The influence of cultural factors on the demography and pattern of gene flow from the Makiritare to the Yanomama Indians. *Am. J. Phys. Anthrop.* **32**, 339.

EWENS, W. J. (1972). The sampling theory of selectively neutral alleles. *Theoret. Pop. Biol.* **3**, 87.

GEERDINK, R. A., BARTSTRA, H. A. & SCHILLHORN VAN VEEN, J. M. (1974). Serum proteins and red cell enzymes in Trio and Wajana Indians from Surinam. *Am. J. Hum. Genet.* (In the Press.)

GERSHOWITZ, H., LAYRISSE, M., LAYRISSE, Z., NEEL, J. V., CHAGNON, N. & AYRES, M. (1972). The genetic structure of a tribal population. II. Eleven blood-group systems and the ABH-Le secretor traits. *Ann. Hum. Genet., Lond.* **35**, 261.

IUCHI, I. (1968). Abnormal hemoglobin in Japan. Biochemical and epidemiologic characters of abnormal hemoglobin in Japan. *Acta Haem. Jap.* **31**, 842.

JOHNSTON, F. E., BLUMBERG, B. S., AGARWAL, S. S., MELARTIN, L. & BURCH, T. (1969). Alloalbuminemia in Southwestern U.S. Indians: Polymorphism of Albumin Naskapi and Albumin Mexico. *Hum. Biol.* **41**, 263.

JOHNSTON, F. E., ALARCON, O., BENEDICT, F., DARY, M., GALBRAITH, M. & GINDHART, P. S. (1973). Albumin
    Mexico (Al^Me) in the Guatemalan Highlands. *Am. J. Phys. Anthrop.* **38**, 27.
MCDERMID, E. M. (1971). Serum albumin variation in Indian populations. *Vox. Sang.* **21**, 462.
MELARTIN, L. & BLUMBERG, B. S. (1966). Albumin Naskapi: a new variant of serum albumin. *Science, N.Y.*
    153, 1664.
MELARTIN, L., BLUMBERG, B. S. & LISKER, R. (1967). Albumin Mexico: a new variant of serum albumin.
    Nature, Lond. **215**, 1288.
NEEL, J. V., ARENDS, T., BREWER, G., CHAGNON, N., GERSHOWITZ, H., LAYRISSE, M., LAYRISSE, Z.,
    MACCLUER, J., MIGLIAZZA, E., OLIVER, W., SALZANO, F., SPIELMAN, R., WARD, R. & WEITKAMP, L. R. (1972).
    Studies on the Yanomama Indians. *Proc. IVth Int. Cong. Human Genetics*, Amsterdam, pp. 96–111.
    Excerpta Medica.
SICK, K., BEALE, D., IRVINE, D., LEHMANN, H., GOODALL, P. & MACDOUGALL, S. (1967). Haemoglobin
    G_{Copenhagen} and Haemoglobin J_{Cambridge}, two new β-chain variants of Haemoglobin A. *Biochim. biophys. Acta*
    140, 231.
SMOUSE, P. E. (1974). Likelihood analysis of geographic variation in allelic frequencies. II. The logit model
    and an extension to multiple loci. *Theoretical and Applied Genetics.* (In the Press.)
SMOUSE, P. E. & KOJIMA, K. (1972). Maximum likelihood analysis of population differences in allelic fre-
    quencies. *Genetics* **72**, 709.
SPIELMAN, R. S., MIGLIAZZA, E. C. & NEEL, J. V. (1974). Linguistic and genetic differences among Yanomama
    language areas. *Science, N.Y.* **184**, 637.
TANIS, R. J., NEEL, J. V., DOVEY, H. & MORROW, M. (1973). The genetic structure of a tribal population,
    the Yanomama Indians. IX. Gene frequencies for 17 serum protein and erythrocyte enzyme systems in
    the Yanomama and five neighbouring tribes; nine new variants. *Am. J. Hum. Genet.* **25**, 655.
UEDA, N., NEEL, J. V., TANIS, R. J., KISHIMOTO, S., FERRELL, R. E. & HAMILTON, H. (1974). Report on
    serum and red blood cell electrophoretic variants among members of the communities of Hiroshima and
    Nagasaki. (In manuscript.)
WARD, R. H. (1972). The genetic structure of a tribal population, the Yanomama Indians. V. Comparisons
    of a series of genetic networks. *Ann. Hum. Genet., Lond.* **36**, 21.
WARD, R. H., GERSHOWITZ, H., LAYRISSE, M. & NEEL, J. V. (1974). The genetic structure of a tribal popula-
    tion, the Yanomama Indians. X. Blood group and secretor differences between the Yanomama and their
    neighbors; the uniqueness of the tribe. (In manuscript.)
WEITKAMP, L. R. & CHAGNON, N. A. (1968). Albumin Maku. A new variant of human serum albumin.
    Nature, Lond. **217**, 759.
WEITKAMP, L. R., FRONGLEN, G., ROKALA, D. A., POLESKY, H. F., SIMPSON, W. E., SUNDERMAN, F. W.,
    BELL, H. E., SAAVE, J., LISKER, R. & BOLS, S. W. (1969). An electrophoretic comparison of human serum
    albumin variants: eight distinguishable types. *Hum. Hered.* **19**, 159.
WEITKAMP, L. R., SALZANO, F. M., NEEL, J., PORTA, F., GEERDINK, R. A. & TARNOKY, A. L. (1973a).
    Human serum albumin: twenty-three genetic variants and their population distribution. *Ann. Hum.
    Genet., Lond.* **36**, 381.
WEITKAMP, L. R., MCDERMID, E. M., NEEL, J. V., FINE, J. M., PETRINI, C., BONAZZI, L., ORTALI, V.,
    PORTA, F., TANIS, R. J., HARRIS, D. J., PETERS, T., RUFFINI, G. & JOHNSTON, E. (1973b). Additional
    data on the population distribution of human serum albumin genes; three new variants. *Ann. Hum.
    Genet., Lond.* **37**, 219.
YANASE, T., HANADA, M., SEITA, M., OHYA, I., OHTA, Y., IMAMURA, T., FUJIMURA, T., KAWASAKI, K. &
    YAMAOKA, K. (1968). Molecular basis of morbidity. *Jap. J. Hum. Genet.* **13**, 40.