Inherited thermostability variants of seven enzymes in a Japanese population

By C. SATOH, J. V. NEEL,* A. MIURA,†, C. UENO, H. ARAKAWA, H. OMINE, K. GORIKI‡ and M. FUJITA

Division of Biochemical Genetics, Department of Clinical Laboratories, Radiation Effects Research Foundation, 5-2 Hijiyama Park, Minami-Ward, Hiroshima 730, Japan

SUMMARY

The frequency of inherited variations in thermostability was investigated in a series of seven enzymes in a Japanese population. Among a total of 5930 determinations, nine variants were encountered. In each instance one parent exhibited a similar finding. It is suggested that this procedure should detect a high proportion of the variants of these enzymes characterized by amino acid substitutions not altering molecular charge. Failure to detect more such thermostability variants is interpreted to mean that electrophoresis not only detects amino acid substitutions altering molecular charge but also a considerable proportion of those that do not alter charge.

INTRODUCTION

The first studies of the genetic consequences of exposure to the atomic bombs were, as necessitated by the state of the art at that time, morphological in nature (Neel & Schull, 1956). Since 1972, however, a major effort to detect mutations affecting protein structure/function has been in progress. In this effort, the application of the technique of electrophoresis to some 30 different proteins has been the principal tool (Neel *et al.* 1980*a*; Satoh *et al.* 1982*a*, *b*), but a subset of these proteins is also being examined for variation characterized by an alteration in the activity or thermostability of the protein. These latter studies are restricted to proteins functioning as enzymes. A report on the frequency of activity variants of 11 enzymes has been recently published (Satoh *et al.* 1983). In the present communication we report on the occurrence of inherited variations in thermostability in a subset of seven of the enzymes studied for the frequency of activity variants. Normative data resulting from 5930 determinations will be presented, after which we will discuss the potential of this approach in a monitoring program. In addition, we will consider the total frequency of variants revealed by a combination of these three approaches in a Japanese population, arguing that electrophoresis is detecting a considerable proportion of so-called 'silent' amino acid substitutions.

MATERIALS AND METHODS

Population. The blood samples used for this study were obtained from children of parents proximally exposed to the atomic bombs in Hiroshima and Nagasaki (one or both parents exposed within 2000 m from hypocentre) and from children of parents distally exposed (one or

Table 1. *Conditions employed in the experiments to determine thermostability or thermolability of enzymes*

| | Heating condition | | | | Activity determination | | | |
| | Temperature | Time (min) | | | Sampling volume | Lag time | Reading points | Reading interval |
Enzyme	(°C)	1	2	3	(μl)	(sec)		(sec)
GOT1	60	10	20	—	10	600	10	60
LDH	60	10	20	—	5	14	10	5
PGK	52·5	20	30	—	5	30	20	5
GPI	52·5	10	20	30	5	120	10	20
PK	52·5	20	30	—	20	30	15	60
6PGD	52·5	20	30	—	20	30	10	30
AK1	50	20	30	—	5	80	30	5

both parents exposed beyond 2500 m from hypocentre). The latter were sex and age matched to the former children. About 30 % of the population is composed of siblings. Further details concerning the two groups of children are given elsewhere (Neel *et al.* 1980*a*; Schull, Otake & Neel, 1981). Inasmuch as all the thermostability variants which were encountered were inherited, the complication of a radiation effect does not enter into this study, and the results from the two sets of children will be combined.

Systems. The seven enzyme systems examined for heat denaturation characteristics were as follows: glutamate-oxaloacetate transaminase-1 (GOT1, E.C. 2.6.1.1), lactate dehydrogenase (LDH, E.C. 1.1.1.27), phosphoglycerate kinase (PGK, E.C. 2.7.2.3), glucose phosphate isomerase (GPI, E.C. 5.3.1.9), pyruvate kinase (PK, E.C. 2.7.1.40) 6-phosphogluconate dehydrogenase (6PGD, E.C. 1.1.1.44) and adenylate kinase-1 (AK1, E.C. 2.7.4.3).

Samples. The samples used in the heat denaturation experiments were the same 1:20 diluted haemolysates as those used in the screening for deficiency variants of these same enzymes. As described previously (Satoh *et al.* 1983), samples were processed within 12–36 h of collection. Leucocytes were excluded by passing whole blood through a column composed of α-cellulose and Sigmacell (R) type 50, and thrice washed packed cells were then either immediately stored in liquid nitrogen (Nitrogen Preserved, NP) or used for analysis within 3 days of collection (FResh samples, FR), depending on the laboratory schedule. Haemolysates were made from 150 μl of packed cells and 2·85 ml of the stabilizing solution described by Beutler (1975) or ICSH (Beutler *et al.* 1977), composed of β-mercaptoethanol (0·7 mM) and EDTA (2·7 mM), pH 7·0.

Procedures. Heat denaturation experiments were performed as follows: (1) Three or four 100 μl aliquots of the haemolysates were prepared, the number depending on the system under study. The first aliquot was maintained in an ice-water bath and served as the initial control. The second to the fourth aliquots were heated in thermistor-regulated water baths set at an appropriate temperature, with ±0·1 °C accuracy, for times as shown in Table 1. (2) Following heating, the tubes were quickly cooled in an ice-water bath and then centrifuged at 3000 r.p.m. (1200 *g*) for 10 min to precipitate denatured proteins. (3) The enzyme activity (i.u./l) remaining in the supernatant of the heated haemolysates and in the untreated haemolysates was determined following the method of Beutler (1975) and ICSH (Beutler *et al.* 1977), using an Aminco Rotochem II a/36 centrifugal analyser, set at 340 nm and 30 °C. The volume of the samples, lag time, number of reading points, and reading intervals are shown in Table 1. (4) The activity

remaining in the supernatant of each of the incubated aliquots was compared with that of the appropriate untreated haemolysate, and the results expressed as percent remaining activity.

As the heat denaturation experiments were conducted on the same day as the enzyme activity measurements screening for deficiency variants, in order to reduce the complexity of the procedure, no special buffer was employed for the heat denaturation experiments, and the conditions were chosen to be applicable to as many enzyme systems as possible. As AK1 was relatively more labile, and GOT1 and LDH relatively more stable, than the other enzymes, they were incubated at 50 and 60 °C, respectively. LDH is a tetramer composed of A- and B-subunits, over 80 % of its activity being derived from the B-subunit. It is accordingly assumed that there is some bias against the detection of thermostability variants of this enzyme and that any variants detected are referrable to the B-subunit.

Data selection. As described in the previous paper on deficiency variants (Satoh *et al.* 1983), a set of operating rules was adopted in order to minimize the contributions of technical error or extraneous biological factors. Based on these rules, certain determinations were excluded not only from the data on enzyme activity but also from the data on heat denaturation. When the activities of duplicate determinations did not agree within 5 %, or when for a given individual the activity values of four or more enzymes were 20 % above or below the mean values for that day, or when on a given day the mean value for a certain enzyme was above or below the cumulative mean by 1 S.D., those values were neither used to calculate the final cumulative mean nor used to select thermostability variants. In addition, in the heat denaturation experiment, one more basis for excluding a result was adopted: if the daily mean of the 'percent remaining activity' of a specific enzyme was above or below the cumulative mean by 2 S.D., or 15 %, at two conditions, the results were deemed spurious and excluded from consideration.

Definition of a thermostability variant. With these procedures and constraints, we arbitrarily define a thermostability variant as one characterized in two independent samples by a percent remaining activity after heating of more than 2·5 S.D. below or above the cumulative mean at two heating temperatures, requiring that at the same time the other six enzymes of the individual exhibiting the variant are of normal stability. This practice guards against false positives as a result of deteriorated samples.

Family studies. Whenever the presence of a variant in a child was confirmed by a repeated determination using a newly obtained sample, an effort was made to analyse samples from both parents and as many siblings as possible, not only with respect to the enzyme in question but for all the enzymes under study. This analysis included the complete battery of electrophoretic and enzyme activity determinations.

<center>RESULTS</center>

Normative values for the heat stability of the seven enzymes studied, including values for both the genetic polymorphisms and the rare electromorphs included in the series, are given in Table 2. It will be apparent that conditions of processing and/or storage affect the apparent thermostability, and that the various electrophoretic variants often have altered thermostabilities (although for some phenotypes the numbers are too small for definite conclusions). All departures from normality (i.e. variants) have been defined both with reference to the conditions under which the sample was processed (FR or NP) and the phenotype. In calculating the

Table 2. *Thermostability profiles of 7 erythrocyte enzymes for fresh (FR) and nitrogen-preserved (NP) samples*

Enzyme	Electrophoretic phenotype	FR				NP			
		No. of determinations	Remaining activity (%)			No. of determinations	Remaining activity (%)		
			Time 1	Time 2 (Mean ± s.d.)	Time 3		Time 1	Time 2 (Mean ± s.d.)	Time 3
GOT1	1	260	78±8	75±8	—	732	82±8	81±10	—
	1-2HR1	9	77±7	75±6	—	33	91±8	91±13	—
	1-3NG1	1	73	64	—	4	86±8	86±14	—
	1-4NG1	3	89±2	86±4	—	6	93±2	94±8	—
	2HR1	0	—	—	—	1	80	86	—
LDH	Normal	251	57±5	42±6	—	376	61±5	46±7	—
PGK	1-Male	113	56±6	46±6	—	101	52±10	42±10	—
	1-Female	135	58±5	47±5	—	112	55±6	44±6	—
GPI	1	341	60±3	38±4	24±3	1031	62±4	40±4	27±4
	1-3HR1	0	—	—	—	4	61±1	39±1	26±2
	1-4HR1	2	50±4	29±4	18±2	2	46±1	25±1	15±0
	1-5HR1	1	62	39	25	0	—	—	—
	1-5NG1	0	—	—	—	2	63±3	40±2	26±2
PK	1	237	56±6	54±6	—	737	52±6	51±6	—
6PGD	A	281	89±4	84±4	—	565	86±5	81±5	—
	AC	63	65±7	59±7	—	114	61±6	54±6	—
	C	5	37±9	30±5	—	8	34±7	27±6	—
	VSA*	0	—	—	—	1	56	51	—
AK1	1	173	43±5	38±5	—	226	37±4	42±4	—

* A heterozygous slow variant with type A.
Description of phenotypes of GOT1 and GPI are shown in *Jpn. J. Hum. Genet.* **26**, 155 (1980) and *Ann. Hum. Genet.* **41**, 419 (1978), respectively.

Table 3. *Number of thermostability determinations and findings in Hiroshima children.*

(Further explanation in text.)

System		No. of determinations		Presumptive variants	False positive	Not yet confirmed	Confirmed and family study positive	Correction factor for variants*1	Frequency of variants/ 1000
GOT1	FR	273	} 1049	5	5	0	0	0·00	0·00
	NP	776							
LDH	FR	251	} 627	6	4	1	1	0·20	1·91
	NP	376							
PGK	FR M	113	} male 214	4	1	0	3	0·00	
	F	135							6·51
	NP M	101	} female 247	0	0	0	0	0·00	
	F	112							
GPI	FR	344	} 1383	7	3	1	3	0·50	2·53
	NP	1039							
PK	FR	237	} 974	1	1	0	0	0·00	0·00
	NP	737							
6PGD	FR	349	} 1037	10	6	3	1	0·43	1·38
	NP	688							
AK1	FR	173	} 399	5	2	2	1	0·67	4·18
	NP	226							
Total	FR	1875	} 5930	38	22	7	9	1·80	2·36*2 (1·82*3)
	NP	4055							

*1 Correction factor for variants is calculated on the assumption that of the 'not yet confirmed', the proportion of true variants would be as the ratio of confirmed: (confirmed + false positive).

*2 Unweighted mean.

*3 Weighted mean.

frequency of thermostability variants and, later, total variant frequencies, we will, to avoid double counting, not score an abnormal electromorph also as a thermostability variant.

Table 3 presents an accounting of the number of *presumptive* variants encountered in the study and the disposition of these variants on further studies. It will be noted that of the 38 apparent variants on the first screening, 22 (58%) were not confirmed on a repeat determination on a second sample, and were excluded from further studies. These 'failures to repeat' with a second sample were usually not borderline but rather well within the normal range, i.e. we do not feel we are identifying a class of 'borderline' thermostability variants. For seven presumptive variants we have been unable to obtain the second sample necessary to confirm the findings with the first. Nine variants have been confirmed by a second sample; family studies (minimum of father and mother) have been carried out on each and in each case one parent or the other exhibited a similar thermodenaturation profile (Table 4). In the last column of Table 3, variant frequencies have been calculated not only on the basis of the confirmed variant(s) but also on the assumption that of the 'not yet confirmed', the proportion of true variants would be as the ratio of confirmed:(confirmed + false positive). Only one of the nine confirmed variants exhibited increased thermostability, so that variants with decreased thermostability are relatively much more common than those with increased thermostability.

There follows a brief discussion of the findings with respect to each system:

GOT1. This series of 1049 determinations included 57 electrophoretic variants, of four types

Table 4. *Family studies of thermostability variants*

Enzyme	Variant MF no.	Family member	Initial activity (i.u./gHb)	Percent remaining activity after heating		
				Time 1	Time 2	Time 3
LDH	Normal		170·5 ± 12·8	61 ± 5	46 ± 7	—
	468309	Child	177·8	80 (3·6)	69 (3·2)	—
		Mother	165·1	68 (1·3)	53 (1·0)	—
		Father	157·7	76 (2·8)	66 (2·8)	—
PGK	1-Male		259·8 ± 15·0	52 ± 010	42 ± 10	—
	1-Female		266·6 ± 15·6	55 ± 6	44 ± 6	—
	415936	Child	252·7	23(−2·9)	12(−3·2)	—
		Mother	278·6	40(−2·6)	28(−2·8)	—
		Father	263·7	54 (0·2)	42 (0)	—
	957756	Child	253·6	8(−4·4)	4(−4)	—
		Mother	228·7	35(−3·5)	25(−3·3)	—
		Father	251·1	59 (0·7)	46 (0·4)	—
	959058	Child	279·2	19(−3·3)	11(−3·3)	—
		Mother	289·9	35(−3·5)	27(−2·9)	—
		Father	—	—	—	—
GPI	1		51·5 ± 4·6	62 ± 4	40 ± 4	27 ± 4
	405707	Child	39·4	42(−5·4)	20(−4·8)	11(−4·1)
		Mother	50·8	65 (0·8)	43 (0·7)	30 (0·8)
		Father	46·3	48(−3·8)	26(−3·3)	15(−3·1)
		Aunt 1	40·5	47(−4·1)	26(−3·3)	15(−3·1)
		Uncle	41·2	45(−4·6)	24(−3·8)	13(−3·6)
		Brother	49·2	65 (0·8)	43 (0·7)	29 (0·8)
		Aunt 2	52·6	64 (0·5)	44 (1·0)	29 (0·5)
	412480	Child	47·2	47(−4·1)	25(−3·6)	14(−3·3)
		Mother	49·5	64 (0·5)	42 (0·5)	29 (0·5)
		Father	47·9	49(−3·5)	27(−3·1)	16(−2·8)
	903626	Child	41·5	47(−4·1)	25(−3·6)	16(−2·8)
		Mother	50·5	46(−4·3)	29(−2·6)	16(−2·8)
		Father	—	—	—	—
6PGD	A		6·0 ± 0·5	86 ± 5	81 ± 5	—
	277473	Child	6·48	58(−5·6)	52(−5·8)	—
		Mother	—	—	—	—
		Father	6·10	50(−7·2)	43(−7·6)	—
AK1	1		202·4 ± 16·9	37 ± 4	32 ± 4	—
	962058	Child	135·2	24(−3·4)	20(−3·1)	—
		Mother	134·1	26(−2·9)	21(−2·8)	—
		Father	208·9	36(−0·2)	30(−0·5)	—

(1) Mean ± s.d. for the normal phenotype (or type 1, or type A) is shown on the first line for each enzyme and data for the family members follow.

(2) Data for children were those obtained from the second new NP samples since all their parents were examined on NP samples and mean ± s.d. shown is for NP samples.

(3) Number in parenthesis shows departure from the mean by s.d. unit.

(Kimura *et al.* 1980). Although the enzymatic activities associated with these variants tend to be low at 30 °C (approximately 80 % of that of GOT1 1) (Satoh *et al.* 1983), the gene products do not appear to be less stable than normal, and one of them (GOT1 4NG1) indeed appears to exhibit an increased thermostability. Otherwise, no abnormalities of thermostability were encountered among the 1049 tests.

LDH. Among a total of 627 determinations, a single variant was encountered, characterized

by increased thermostability. The father of this girl showed similar findings; the mother was normal (Table 4). This is the only example of a hereditary variant with increased thermal stability encountered in the 5930 determinations. Electrophoretic patterns of the proposita and her father suggested that the variant involved the A-subunit, since the intensity of the LDH3 and LDH4 bands was weak. Among the 3121 samples subjected to electrophoretic screening between October 15 1979 and September 28 1981, nine additional cases of similar A-subunit deficiency were encountered. Family studies could be performed on seven of these nine cases. Two of these seven had been included in the thermostability series and found to be within normal limits, as were their parents, but both the affected parents and children had values 2 s.d. above normal. In the other five cases, the values for the propositi (NP samples) were 75 ± 5 and $62 \pm 6 \%$ after heating for 10 and 20 min, respectively. In each case, one parent showed the electrophoretic A-subunit deficiency. The thermostability values were 75 ± 3 and $62 \pm 4 \%$ for the five affected parents and 66 ± 3 and $50 \pm 4 \%$ for the five normal parents. The values of the affected children and parents are elevated values but technically do not meet our criteria for a thermostability variant. We conclude that we have identified among the above-mentioned electrophoretic series of 3121 determinations at least one and possibly several deficiencies of the A-subunit of LDH, in consequence of which the remaining B-subunits exhibit increased thermostability, although in general not to a degree qualifying for the term 'thermostability' as used in this paper.

PGK. This enzyme is the only sex-linked trait in the series. The thermostability profiles of males and females are identical. Among the 214 determinations on males, there were three exhibiting increased lability (Table 3). In each instance, as expected with a sex-linked trait, the mother's enzyme exhibited increased thermolability but the father's was normal. No thermostability variants were encountered among the 247 females who were screened.

We have previously reported that this series includes three male siblings exhibiting significantly decreased enzyme activity at 30 °C (approximately 40 % of normal) (Satoh *et al.* 1983). The enzyme of these three males (Master File nos. 958 160, 958 161, 958 162) also exhibited increased thermolability, as did that of the mother. Since, however, these variants have already been included in a tabulation of activity variants, they are not included in Table 3.

GPI. Four different types of rare electrophoretic variants of this enzyme (Tanis *et al.* 1978) were represented in the sample of 1383 examinations. As shown in Table 2, 3 of these exhibited normal thermostability but all four individuals with the 1-4HR1 phenotype exhibited increased lability, of about the same degree, confirming the earlier observations of Satoh & Mohrenweiser (1979) concerning the thermostabilities of electrophoretic variants of this enzyme.

Three thermolability variants have been encountered in this enzyme series, all confirmed by family studies. In one instance (Table 4), in which the father was also affected, the family studies were extended to include a brother, two paternal aunts, and one paternal uncle; the uncle and one of the aunts were also trait carriers. This particular variant exhibited 77 % of normal activity at 30 °C, and is mentioned in our previous paper on activity variants (Satoh *et al.* 1983), but not scored as an activity variant. With respect to the other two variants in both cases one parent exhibited the trait, so that the genetic nature of the variant is confirmed even though the father of one child refused examination.

PK. No thermostability variants of this enzyme were detected in the course of 974 determinations. This seems surprising in view of the high frequency of variants of this enzyme

with low activity (13·8/1000 determinations; Satoh *et al.* 1983). This series included 9 persons heterozygous for a low activity variant; none of these showed increased thermolability, a finding to be expected since it is the thermolability of the product of the normal allele which is being measured.

6PGD. In contrast to the enzymes discussed thus far, this system is characterized by a relatively common polymorphism, so that the 1037 determinations include homozygotes as well as heterozygotes for this (C) variant. The heterozygote (type AC) shows marked thermolability, and the homozygote (type C), very marked lability. Within the limits of error, the loss in the homozygote is twice that in the heterozygote. The activities of these three phenotypes at 30 °C (NP samples) differed only slightly (A: 6·0±0·5 i.u./gHb; AC: 5·6±0·5 i.u./gHb; C: 5·4±0·6 i.u./gHb).

Aside from the above, two thermolability variants were encountered. The loss of activity in the first variant (MF no. 277473) was similar to that in type AC samples though it was type A and its activity was normal (6·48 i.u./gHb) for its type. His father who was type A also showed the same degree of thermolability (Table 4). The second variant (MF no. 241070) had already been identified as a slowly-migrating enzyme band, and so in our treatment of variant frequencies (below) will be tabulated with the electromorphs.

AK1. A single thermolability variant was encountered in 399 determinations. This individual had previously been noted to have a low activity at 30 °C, but the percent normality (67 %) was one percentage point above our arbitrary cut-off level for activity variants (Satoh *et al.* 1983). The carrier mother of this person also exhibited a low initial activity (66 % of normal). Rather arbitrarily, we score this as a thermostability variant.

Although we do not feel the data are sufficiently extensive to permit across-locus comparisons as to types of variants, we note that less genetic variation has been encountered with respect to this enzyme than any other so far studied: no electrophoretic variant in 10817 electrophoretic examinations, and no activity variants in 2863 determinations other than the borderline case mentioned above (which fully qualifies as a thermostability variant).

DISCUSSION

This study confirms the demonstration of Mohrenweiser & Neel (1981) that, under survey conditions, it is possible to identify a class of inherited biochemical variants solely by alterations in thermostability. Their frequency in this study of 0·0018, with 95 % confidence limits of 0·0008 and 0·003, is in agreement with the frequency of 0·0038, with 95 % confidence limits of 0·0008 and 0·0110, encountered in the predominantly Caucasoid population in Ann Arbor, Michigan. For reasons detailed earlier, this is thought to be a minimal estimate of the frequency of such variants. Given the invariable occurrence in one parent of a finding similar to that in the proband, a finding lying outside the normal distribution, we conclude we are recognizing a qualitative trait exhibiting monogenic inheritance.

Six enzymes have now been studied in Japanese and Caucasians for the frequency of three kinds of variants, namely, electrophoretic, activity, and thermostability variants. The results are summarized in Table 5. For these purposes, an activity variant has been arbitrarily defined as one which in a heterozygote results in enzyme levels 66 % of normal. The basis and empirical justification for that designation has been discussed in detail elsewhere (Mohrenweiser, 1981).

Table 5. *Frequency of 'rare' variants of six enzymes per* 1000 *determinations in a Japanese population (Hiroshima and Nagasaki) and in a U.S. predominantly Caucasoid population.*

(The enzymes are GOT1, LDH, PGK, GPI, PK and AK1.)

Type of variant	Group	Number of determinations	Number of variants	Frequency/ 1000	95 % confidence interval	Reference
Electrophoretic[*1]	Japanese	38087	117	3·07	2·5–3·6	Neel *et al.* 1980*a*
	Caucasoid	8484	21	2·47	1·5–3·8	Neel *et al.* 1980*b*
Activity	Japanese	18232	54(60·8)[*2]	3·33[*3]	2·5–4·2	Satoh *et al.* 1983
	Caucasoid	6428	6	0·93	0·3–2·0	Mohrenweiser, 1981, 1983
Thermostability	Japanese	4893	8(9·37)[*2]	1·91[*4]	0·8–3·5	This paper
	Caucasoid	500[*5]	2	4·00	0·5–14·4	Mohrenweiser & Neel (1981)

[*1] Electrophoretic screening for PK and PGK were not performed. Numbers of determinations and variants were based on the other four enzymes. The frequency/1000 is a mean for the four enzymes.

[*2] Numbers of variants in parentheses are calculated on the same assumption as was used to calculate the frequency of variants/1000 in Table 3 of this paper.

[*3] Mean calculated with the frequencies in Table 5 of Satoh *et al.* (1983).

[*4] Mean calculated with the frequencies in Table 3 of this paper.

[*5] GOT1 not screened for Caucasoid.

There is potential overlap in this classification, i.e. there are many examples of an electrophoretic variant also exhibiting reduced activity and/or decreased thermostability (Spencer, Hopkinson & Harris, 1964; Yoshida, Beutler & Motulsky, 1971; Milkman, 1976); to prevent 'double counting', precedence in classification is given to electromorphs (regardless of other properties), then to activity variants, and finally to thermostability variants. We have limited the Caucasoid sample of Table 5 to a predominantly Caucasoid population from Ann Arbor, Michigan, because the techniques for detecting these three types of variants were introduced into the Japanese study from Ann Arbor, thus ensuring maximum comparability. The technique for determining thermostability variants, however, differed slightly in the two laboratories, the Ann Arbor study holding the samples at each of three different temperatures for 20 min. The simpler approach adopted in the present study seems more appropriate to large-volume screening. (Failure of these frequencies per 1000 to agree precisely with those published earlier is due to the fact that here we express the results with a weighted mean.)

Previous studies (Ferrell *et al.* 1977; Ueda *et al.* 1977; Satoh *et al.* 1977; Tanis *et al.* 1978) have established that among the 22 proteins then being studied in this laboratory, the frequency of electrophoretic variants in the enzymes *not* studied for activity is somewhat higher than in those which were so studied. If, as seems probable, there is intra-locus correlation in the frequencies of these three types of variants, then our estimates of the frequency of activity and thermostability variants could be on the low side.

The data on activity and thermostability variants have not yet reached the point where specific locus-type comparisons across ethnic groups seem indicated. On the other hand, the data are sufficient for a comparison by type of variant. Because of the smallness of some of the numerators, we have calculated confidence intervals based on the expectation of a Poisson variable for numerators below 50. With this procedure, there appears to be a significant difference between the two populations only with respect to the frequency of activity variants. An analysis of Table 5 for heterogeneity, which would usually be the next step, is precluded

by the overlapping nature of the samples on which the frequencies are based. The total frequency of these three types of variants in the two populations is quite similar, 8·3/1000 in the Japanese and 7·4/1000 in the US Caucasoids.

These data bear importantly on the issue of average locus heterozygosity in human populations. In estimating this heterozygosity, it has often been assumed that electrophoresis detects only variants in which an amino acid substitution has altered molecular charge. Since it can be readily calculated that only about one third of mutations resulting in an amino acid substitution will alter molecular charge, and since selection against mutations resulting in uncharged amino acid substitutions should not be greater than against charged, and possibly less, it is generally assumed there is a large pool of 'silent' genetic variation awaiting discovery. The present study raises questions as to the size of this pool. As we have noted, the enzymes examined in this study are electrophoretically less variable than most proteins examined by one dimensional electrophoresis. Thus, whereas the average Index of Heterozygosity for a series of 23 proteins in Japanese (including those covered by this report; Neel *et al.* 1978, table 2) was 7·7 %, for these seven enzymes the Index is 3·3 %. Generalization from this series to the average should proceed cautiously. Of the 245 heterozygotes for electromorphs included in Table 2, 196 (80 %) exhibited increased thermolability. Thus assuming electrophoresis detects only charged substitutions, thermostability studies would have detected 80 % of the charged amino acid substitutions. This small series, however, is dominated by the C variant of 6PGD. It is an interesting coincidence that in a review of the literature on experimental organisms we have suggested that 'at least two thirds and possibly three quarters of the variants recognized by electrophoresis exhibit easily detected thermolability...' (Neel, 1984). While it is possible that charged amino acid substitutions are more likely to alter thermostability than uncharged, let us assume this is not so. Then our thermostability studies should be detecting the majority of the uncharged substitutions not detected by electrophoresis. This increment, of 0·2 %, is a much smaller increment than anticipated. At least a partial explanation may be found in the growing body of opinion that electrophoresis is also detecting some portion of the uncharged amino acid substitutions, presumably because of conformational changes in the molecule (cf. Bonhomme & Selander, 1978; Ayala, 1982). The pool of genetic variation characterized by 'silent' amino acid substitutions which was presumed to be awaiting discovery is almost certainly substantially less than projected from the 1 charged : 2 uncharged ratio mentioned earlier.

Elsewhere (Neel, 1984) we have discussed the paradox that as estimates of protein variability are being revised downwards, not only because of results such as those presented by the present study but because of studies with two-dimensional polyacrylamide gel electrophoresis, estimates of variability at the DNA level are escalating rapidly. We suggested that the resolution is to be found in a very non-random distribution of DNA restriction fragment length polymorphisms, for which there is already some evidence. The next decade will witness rapid progress in this arena of inquiry, with its associated implications regarding the relative roles of mutation and selection in the origin of this non-randomness. As probes for more and more structural genes are developed, this decade will also witness the ability to confirm at the DNA level not only the correctness of the conclusions from the analysis of Table 5, but, in addition, to test to what extent the numerical similarities are the result of the same or of different mutations.

With reference to the usefulness of thermostability variants in a study of mutation rates, we

estimate that the effort involved in identifying a thermostability variant (including exclusion of false positives as well as routine laboratory procedures) is in our hands approximately four times that involved in detecting an electromorph, and two times that involved in detecting an activity variant. Given the evidence above, that electrophoresis is detecting a higher proportion of all amino acid substitutions than usually assumed, we conclude that this approach is substantially less efficient than the search for mutations altering electrophoretic or activity characteristics.

REFERENCES

AYALA, F. J. (1982). Genetic variation on natural populations: problem of electrophoretically cryptic alleles. *Proc. Natl. Acad. Sci. USA* **79**, 550–554.

BEUTLER, E. (1975). *Red Cell Metabolism. A Mannual of Biochemical Methods*, 2nd ed. New York, San Francisco, London: Grune & Stratton.

BEUTLER, E., BLUME, K. G., KAPLAN, J. C., LOHR, G. W., RAMOT, B. & VALENETINE, W. N. (1977). International Committee for Standardization in Hematology: Recommended methods for red-cell enzyme analysis. *Brit. J. Haematol.* **35**, 331–340.

BONHOMME, F. & SELANDER, R. K. (1978). Estimating total genetic diversity in the house mouse. *Biochem. Genet.* **16**, 287–297.

FERRELL, R. E., UEDA, N., SATOH, C., TANIS, R. J., NEEL, J. V., HAMILTON, H. B., INAMIZU, T. & BABA, K. (1977). The frequency in Japanese of genetic variants of 22 proteins. I. Albumin, ceruloplasmin, haptoglobin, and transferrin. *Ann. Hum. Genet.* **40**, 407–418.

KIMURA, Y., KANEKO, K., MASUNARI, N., TAKAHASHI, N., FUJITA, M., HAZAMA, R., GORIKI, K. & SATOH, C. (1980). Genetic variants of erythrocyte glutamate-oxaloacetate transaminase (soluble GOT) in Japanese populations (Hiroshima & Nagasaki). *Jpn. J. Hum. Genet.* **26**, 155.

MILKMAN, R. (1976). Further evidence of thermostability variation within electrophoretic mobility classes of enzymes. *Biochem. Genet.* **14**, 383–387.

MOHRENWEISER, H. W. (1981). Frequency of enzyme deficiency variants in erythrocytes of newborn infants. *Proc. Natl. Acad. Sci. USA* **78**, 5046–5050.

MOHRENWEISER, H. W. & NEEL, J. V. (1981). Frequency of thermostability variants: estimation of total 'rare' variant frequency in human populations. *Proc. Natl. Acad. Sci. USA.* **78**, 5729–5733.

MOHRENWEISER, H. W. (1983). Enzyme deficiency variants: frequency and potential significance in human populations. *Isozymes, Curr. Top. Biol. Med. Res.* **10**, 51–68.

NEEL, J. V. & SCHULL, W. J. (1956). *The Effect of Exposure to the Atomic Bombs on Pregnancy Termination in Hiroshima and Nagasaki.* Washington, D.C.: NAS-NRC Publ. no. 461.

NEEL, J. V., UEDA, N., SATOH, C., FERRELL, R. E., TANIS, R. J. & HAMILTON, H. B. (1978). The frequency in Japanese of genetic variants of 22 proteins. V. Summary and comparison with data on Caucasians from the British Isles. *Ann. Hum. Genet.* **41**, 429–441.

NEEL, J. V., SATOH, C., HAMILTON, H. B., OTAKE, M., GORIKI, K., KAGEOKA, T., FUJITA, M., NERIISHI, S. & ASAKAWA, J. (1980a). Search for mutations affecting protein structure in children of atomic bomb survivors: Preliminary report. *Proc. Natl. Acad. Sci. USA* **77**, 4221–4225.

NEEL, J. V., MOHRENWEISER, H. W. & MEISLER, M. H. (1980b). Rate of spontaneous mutation at human loci encoding protein structure. *Proc. Natl. Acad. Sci. USA* **77**, 6037–6041.

NEEL, J. V. (1984). A revised estimate of the amount of genetic variation in human proteins: implications for the distribution of DNA polymorphisms. *Am. J. Hum. Genet.* (In the Press.)

SATOH, C., FERRELL, R. E., TANIS, R. J., UEDA, N., KISHIMOTO, S., NEEL, J. V., HAMILTON, H. B. & BABA, K. (1977). The frequency in Japanese of genetic variants of 22 proteins. III. Phosphoglucomutase-1, phosphoglucomutase-2, 6-phosphogluconate dehydrogenase, adenylate kinase, and adenosine deaminase. *Ann. Hum. Genet.* **41**, 169–183.

SATOH, C. & MORENWEISER, H. W. (1979). Genetic heterogeneity within an electrophoretic phenotype of phosphoglucose isomerase in a Japanese population. *Ann. Hum. Genet.* **42**, 283–292.

SATOH, C., AWA, A. A., NEEL, J. V., SCHULL, W. J., KATO, H., HAMILTON, H. B., OTAKE, M. & GORIKI, K. (1982a). Genetic effects of atomic bombs. In *Human Genetics*, Part A, *The Unfolding Genome*, Proceeding of the Sixth International Congress of Human Genetics, 1981, Jerusalem (ed. B. Bonné-Tamir, T. Cohen and R. M. Goodman), pp. 267–276. New York: Alan R. Liss.

SATOH, C., GORIKI, K., HAMILTON, H. B. & NEEL, J. V. (1982b). Search for mutations affecting protein structure in children of atomic bomb survivors. *Jpn. J. Genet.* **57**, 702.

SATOH, C., NEEL, J. V., YAMASHITA, A., GORIKI, K., FUJITA, M. & HAMILTON, H. B. (1983). The frequency among Japanese of heterozygotes for deficiency variants of 11 enzymes. *Am. J. Hum. Genet.* **35**, 656–674.

Schull, W. J., Otake, M. & Neel, J. V. (1981). Genetic effects of the atomic bombs: reappraisal. *Science* **213**, 1220–1227.

Spencer, N., Hopkinson, D. A. & Harris, H. (1964). Quantitative differences and gene dosage in the human red cell acid phosphatase polymorphism. *Nature* **201**, 299–300.

Tanis, R. J., Ueda, N., Satoh, C., Ferrell, R. E., Kishimoto, S., Neel, J. V., Hamilton, H. B. & Ohno, N. (1978). The frequency in Japanese of genetic variants of 22 proteins. IV. Acid phosphatase, NADP-isocitrate dehydrogenase, peptidase A, peptidase B and phosphohexose isomerase. *Ann. Hum. Genet.* **41**, 419–428.

Ueda, N., Satoh, C., Tanis, R. J., Ferrell, R. E., Kishimoto, S., Neel, J. V., Hamilton, H. B. & Baba, K. (1977). The frequency in Japanese of genetic variants of 22 proteins. II. Carbonic anhydrase I and II, lactate dehydrogenase, malate dehydrogenase, nucleoside phosphorylase, triose phosphate isomerase, haemoglobin A and haemoglobin A_2. *Ann. Hum. Genet.* **41**, 43–52.

Yoshida, A., Beutler, E., Motulsky, A. G. (1971). Table of human glucose-6-phosphate dehydrogenase variants. *Bull. W.H.O.* **45**, 243–253.