

*Original investigations***Heterozygosity and ethnic variation in Japanese platelet proteins**Jun-ichi Asakawa¹, James V. Neel², Norio Takahashi¹, Chiyoko Satoh¹, Satomichi Kaneoka¹, Eiko Nishikori¹, and Mikio Fujita¹¹Division of Biochemical Genetics, Department of Genetics, Radiation Effects Research Foundation, Hiroshima, 732 Japan²Department of Human Genetics, University of Michigan Medical School, Ann Arbor, MI 48109, USA

Summary. Sixty-two polypeptides visualized on silver-stained two-dimensional polyacrylamide gel electrophoresis (2-D PAGE) preparations of platelets from 85 Japanese subjects, in total 5,252 spots, have been scored for genetic variation. Inherited variants of 11 of the polypeptides were recognized; the index of heterozygosity was $2.4\% \pm 0.2\%$. Thus far, 10 genetic polymorphisms identified by 2-D PAGE of plasma, erythrocytes, or platelets have been identified in both Japanese and Caucasian subjects. A comparison of allele frequencies reveals four significant ethnic differences. We also observed four polypeptides exhibiting a low frequency polymorphism in one group but not in the other, as well as three polymorphisms in Caucasians for which no counterpart polypeptide has thus far been recognized in the Japanese group and, vice versa, 11 such polymorphisms in Japanese. Although a similar comparison of 7 enzyme polymorphisms studied with one-dimensional electrophoresis (1-D E) in the same populations revealed a relatively higher number of significant ethnic differences, evidence is presented that this is due primarily to the greater number of 1-D E observations entering into the calculation. It is argued that this similarity in the frequency of ethnic differences among the polypeptides studied by 2-D PAGE and by 1-D E is further evidence that the proteins revealed by 2-D PAGE do not differ greatly in their response to the interplay of mutation, selection, and drift from those revealed by 1-D E studies of plasma proteins and erythrocyte enzymes.

allele frequencies at what appear to be identical loci in Caucasians residing in Ann Arbor, Michigan and Japanese from Hiroshima or Nagasaki, and compare the magnitude of these differences with the differences revealed by studies applying the more familiar technique of one-dimensional electrophoresis (1-D E) to erythrocyte enzymes from these same two populations. Some apparent identities in the polymorphisms of platelets and lymphocytes will also be discussed.

Materials and methods

The 85 mother-father-child trios of platelet preparations included in this study were derived from blood samples collected in the course of a study on the genetic effects of the atomic bombs (Neel et al. 1980). In each instance a rare plasma protein or erythrocyte isozyme variant had been encountered in a child in the study group and blood samples had then been obtained from the child's father and mother to determine whether the variant was inherited or the result of a mutational event in one of the parents. At the time of the family studies, a repeat sample had also been obtained from the child.

Platelets were isolated from platelet-rich plasma by centrifugation and washed three times with phosphate-buffered saline. Cell pellets were solubilized by adding two volumes of a solution containing 9 M urea, 2% Nonidet P-40, 0.8% Ampholine (LKB, pH 3.5–10.0), 2% 2-mercaptoethanol, and 1 mM phenylmethylsulfonyl fluoride. Solubilized samples were centrifuged in a microfuge for 2 min and were stored at -80°C until used.

Two-dimensional electrophoresis was carried out using the Iso-Dalt system (Electro-Nucleonics, Oak Ridge, TN) as described by Neel et al. (1984). A 7- μl sample of the supernatant of the solubilized platelet protein preparation was applied to the first-dimension gel. Isoelectric focusing was done at 800 V for 16 h and 1200 V for 2 h in the presence of a 9 M urea and 2.0% Ampholine mixture (0.8% pH 3.5–10.0, 0.8% pH 5–7, and 0.4% pH 5–8). The first-dimension gels were run in a constant-temperature room at 25°C . Second-dimension electrophoresis was performed as previously described (Asakawa et al. 1985). The polypeptides in the gels were visualized by the simplified silver-staining method of Morissey (1981). In this procedure the usual glutaraldehyde step is omitted and the duration of exposure to the dithiothreitol solution was prolonged to 30 min. Hemoglobin charge standards, described by Anderson and Hickman (1979), were prepared as previously (Asa-

Introduction

Recently we have been systematically exploring the kind and amount of genetic variation revealed by two-dimensional polyacrylamide gel electrophoresis (2-D PAGE) of various components of human blood, as a necessary preliminary to decisions concerning the usefulness of this technique in the study of mutation rates, ethnic classifications, human evolution, etc. Since, indirectly, the technique scans some hundreds of kilobases of exon DNA, it also has the potential of identifying functional genetic loci of unusual interest, which with the appropriate techniques could become available for detailed molecular analysis. We have now extended our studies to the platelets of Japanese residing in Hiroshima. Following the presentation of these data we will review all the studies to date employing 2-D PAGE, with reference to ethnic differences in

kawa and Satoh 1986). Molecular weight standards were obtained from BIO-RAD.

A preliminary set of 82 spots was selected for scoring on the basis of our standard criteria of relative isolation from other spots and a staining intensity such that a variant spot de-

rived from any one of them, which presumably would usually exhibit about half the normal staining intensity, should be clearly visible. A further criterion was that in a series of gels, at least 95% of the representatives of any particular spot could be unambiguously scored. As the scoring progressed, this re-

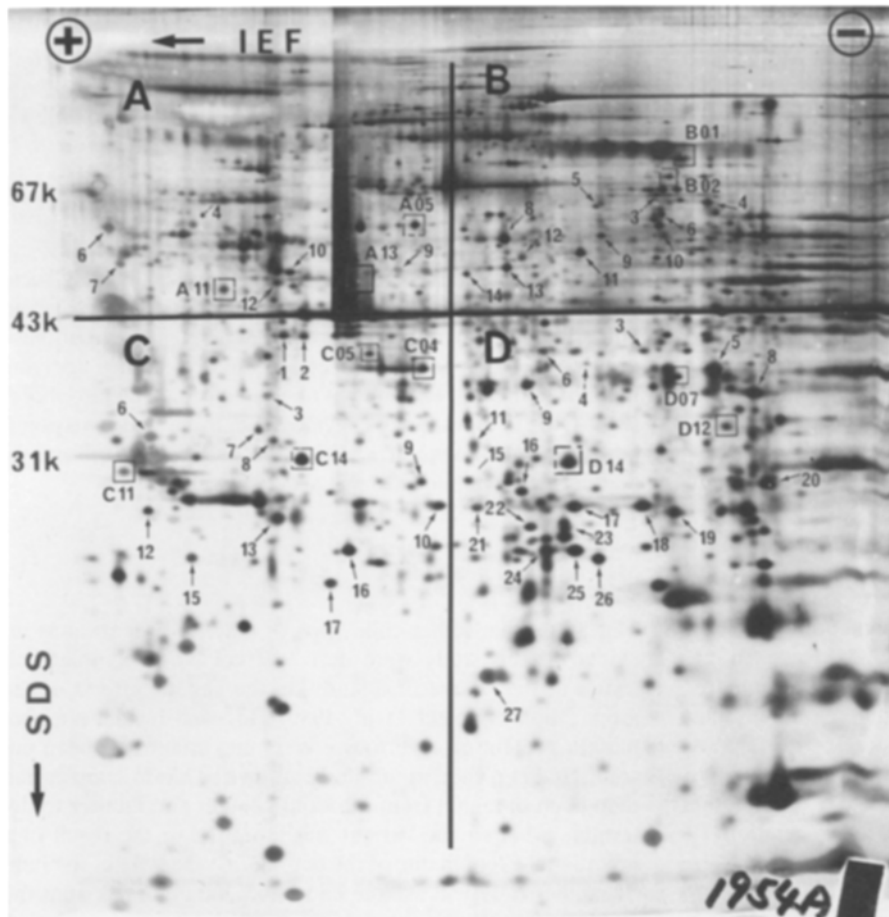


Fig. 1. Silver stained 2-D PAGE pattern of platelet proteins. The gel pattern has been subdivided into four convenient regions (A-D), and the polypeptides suitable for scoring in each region numbered consecutively. The variable polypeptides are surrounded by squares

Table 1. Characteristics of the variant polypeptides encountered in this study

Spot number	~ Mol. wt.	Shift in isoelectric focusing axis	Calibrated charge change	Illustrated in Fig.	Phenotype frequencies				Allele frequencies		
					N	NV	V	Total	<i>p</i>	<i>q_A</i>	<i>q_B</i>
A-005	59,000	V _A -2 mm V _B +3 mm	-0.19 +0.23	2a 2a	78	5	0	85	0.959	0.029	0.012
A-011	47,000	+2 mm	+0.34	2b	84	1	0	85	0.994	-	0.006
A-013	48,000	+3 mm	+0.20	2c	81	4	0	85	0.976	-	0.024
B-001 (transferrin)	75,000	see text	see text	-	72	13	0	85	see text	see text	see text
B-002	69,000	+7 mm	+0.45	2d	55	25	5	85	0.794	-	0.206
C-004	38,000	-4 mm	-0.32	2e	80	5	0	85	0.970	0.030	-
C-005	39,500	+3.5 mm	+0.23	2f	76	8	1	85	0.940	-	0.060
C-011	30,000	+1.5 mm	off-scale	2g	81	4	0	85	0.976	-	0.024
C-014 ^a	31,000	+9 mm	+0.64	2h	72	12	1	85	0.918	-	0.082
D-007	38,000	+1.5 mm	+0.15	2i	82	3	0	85	0.982	-	0.018
D-012 (esterase D)	34,000	-8 mm	-0.85	2j	34	35	16	85	0.606	0.394	-
D-014	30,500	+10 mm	+1.40	2k	59	24	2	85	0.835	0.165	-

^a This variant also is downwardly displaced in the molecular weight axis by about 1 mm (see Fig. 2h)

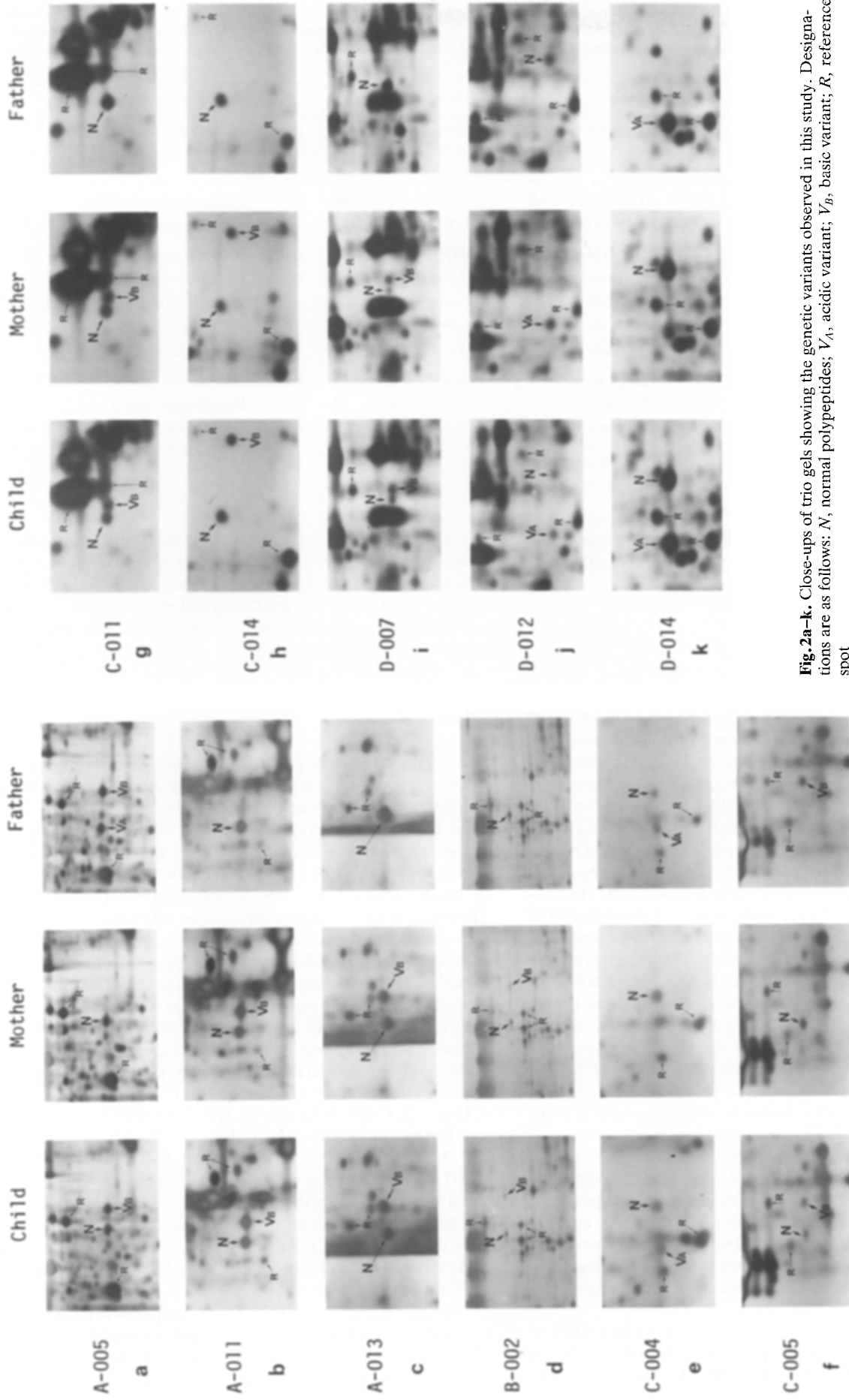


Fig. 2a-k. Close-ups of trio gels showing the genetic variants observed in this study. Designations are as follows: N, normal polypeptides; V_A , acidic variant; V_B , basic variant; R, reference spot

Table 2. A 2×2 χ^2 contrast of allele frequencies for the 10 polypeptides identified in 2-D PAGE preparations that are polymorphic in both Japanese and Caucasians. N designates the number of alleles whose products were scored, n the number of normal alleles, and p the frequency of the normal allele. The identities of the spots where known are: plasma C-014; apolipoprotein E, erythrocyte C-022; glyoxalase 1 and platelet PL-4 (D-012); esterase D

Material	Japanese					Caucasian					χ^2
	Spot designation	N	n	p	Reference	Spot designation	N	n	p	Reference	
Plasma	C-014	220	196	0.89	Asakawa et al. 1985	C-014	124	100	0.81	Rosenblum et al. 1983	4.71 ^a
	D-005	218	116	0.53	Asakawa et al. 1985	D-005	108	62	0.57	Rosenblum et al. 1983	0.51
Erythrocyte	A-003	194	188	0.97	Takahashi et al. 1986	B-011	48	44	0.92	Rosenblum et al. 1985	2.67
	C-009	200	175	0.88	Takahashi et al. 1986	B-009	52	45	0.87	Rosenblum et al. 1985	0.03
	C-022	200	182	0.91	Takahashi et al. 1986	D-014	50	25	0.50	Rosenblum et al. 1985	47.21 ^b
	D-024	200	179	0.90	Takahashi et al. 1986	E-013	52	36	0.69	Rosenblum et al. 1985	13.54 ^b
	D-027	200	146	0.73	Takahashi et al. 1986	D-013	52	40	0.77	Rosenblum et al. 1985	0.33
Platelet	C-005	170	160	0.94	This paper	PLA 2	88	77	0.88	Hanash et al. 1986a	3.40
	C-014	170	156	0.92	This paper	PL-5	88	86	0.98	Hanash et al. 1986a	3.54
	D-012	170	103	0.61	This paper	PL-4	88	71	0.81	Hanash et al. 1986a	10.66 ^b

^a $0.01 < P < 0.05$

^b $P < 0.001$

duced the original battery of 82 spots to 63. Spot designations are based on area of occurrence in the gel and a sequential numbering system.

Results

The polypeptides included in the final battery are designated in Fig. 1. Genetic variation was encountered in 12 of the 63 polypeptides selected for final scoring. At the time the scoring for this study was initiated, we did not know the identity of any of the spots being scored. However, 16 of the blood samples on which this study is based had been collected for verification of a serum transferrin variant detected with 1-D starch gel electrophoresis. When in 13 preparations the phenotype of B-001 varied concurrently with the known presence of a plasma transferrin variant, B-001 was assumed to be transferrin. Whether it should be regarded as a normal constituent of platelet cytosol or functioning membrane, or as a contaminant, is not clear. The three transferrin variants that were not detected with 2-D PAGE had already been found not to be distinguishable from normal with isoelectric focusing (Fujita et al. 1985). Spot B-001 must accordingly be excluded from the battery of polypeptides scored, since its inclusion would introduce a bias into the calculation of the index of heterozygosity; this reduces the number of proteins scored to 62, of which 11 exhibited genetic variation, as summarized in Table 1 and pictured in Fig. 2. The phenotypes for the variable spots were all in Hardy-Weinberg equilibrium.

Polypeptide D-012 was observed in the course of the study to vary in synchrony with the results of 1-D E typings for esterase D on these same samples, and is assumed to be esterase D. Since none of the samples studied had been selected because of the occurrence of an esterase D variant, the results of classifying this spot for variation have been allowed to remain in the series.

Our spot D-014 appears to be identical to spot PL-3 of Hanash et al. (1986a). PL-3 was found to be polymorphic in Caucasians and a similar variant was encountered in this

series, but the most frequent gene product in the Japanese series (which appears to correspond to a similar product in Caucasoid lymphocytes; see below) appears to have been designated as "variant" in the Caucasoid platelet series. Until this discrepancy can be clarified, we will not include this variant in our ethnic comparisons.

Discussion

Among the 62 platelet polypeptides selected in as unbiased a fashion as possible for scoring for genetic variability, 11 (18%) exhibited such variation; the index of heterozygosity for the total material by direct count was $2.4\% \pm 0.2\%$ (128/5252). In similar studies on Caucasian platelets, only 33 spots were scored, of which 3 (9%) exhibited genetic variation; the index of heterozygosity was $2.6\% \pm 0.4\%$ (Hanash et al. 1986a). However, in the Caucasian material an additional 3 spots not included in the "official" battery were also observed to be polymorphic. As is well known, although in any one laboratory with well-standardized techniques the gel patterns are quite reproducible, even minor differences in technique between laboratories can render interlaboratory comparisons difficult. In this instance, among the total of 36 spots scored in Ann Arbor, we feel reasonably secure only in the following 18 identities (Ann Arbor designation first): 2, C-009; 3, D-014; 4, D-012; 5, C-014; 9, D-013; 10, C-012; 11, C-013; 13, D-016; 15, D-019; 16, C-015; 18, C-016; 20, D-022; 21, D-023; 22, D-025; 23, D-026; 25, C-017; 32, D-027, and PLA 2, C-005. Thus, about half the polypeptides scored in Ann Arbor could be recognized with confidence on the Japanese gels, on the basis of the constellation of polypeptides of which they were a part and their relative intensities. These identities include three of the six polypeptides found to be polymorphic in the Caucasian series, as shown in Table 2.

Previous studies by ourselves of polypeptides selected only for their suitability for scoring (i.e., with no prior knowledge of their variability) have revealed heterozygosity indices as follows: Caucasian plasma, $6.2\% \pm 0.7\%$ (Rosenblum et al.

Table 3. A χ^2 contrast of allele frequencies for 7 polymorphic systems of erythrocyte enzymes studied in both Caucasians and Japanese with 2-D E in our laboratories. *T* designates the number of alleles whose products were scored, *N* the number of normal alleles, and *p* the frequency of the normal allele. Data on the Caucasian material will be found in Mohrenweiser et al. (in press) and on the Japanese material in J. V. Neel, C. Satoh, J. Asakawa, N. Takahashi, M. Fujita, and K. Goriki (in preparation)

Enzyme	Adjustment factor	Caucasian obs.	Japanese		Total ^a	χ^2
			Obs.	Adj.		
Phosphoglucomutase 1	0.807	<i>N</i> 4352	(35293)	28471	32823	0.847
		<i>T</i> 5654	(46190)	37262	42916	
		<i>p</i> 0.770		0.764		
Acid phosphatase 1	0.808	<i>N</i> 3429	(35953)	29045	32474	941.8 ^b
		<i>T</i> 5650	(45384)	36664	42314	
		<i>p</i> 0.607		0.792		
Glutamate-pyruvate transaminase	0.832	<i>N</i> 1246	(19941)	16597	17843	28.8 ^b
		<i>T</i> 2324	(33616)	27979	30303	
		<i>p</i> 0.536		0.593		
Adenosine deaminase	0.806	<i>N</i> 5352	(45952)	37034	42386	54.6 ^b
		<i>T</i> 5614	(47296)	38117	43731	
		<i>p</i> 0.953		0.972		
Adenylate kinase 1	0.807	<i>N</i> 5374	(46227)	37323	42697	1519.2 ^b
		<i>T</i> 5604	(46230)	37325	42929	
		<i>p</i> 0.959		1.000		
6-phosphogluconate-dehydrogenase	0.806	<i>N</i> 5471	(42642)	34372	39843	320.1 ^b
		<i>T</i> 5612	(47262)	38096	43708	
		<i>p</i> 0.975		0.902		
Esterase D	0.806	<i>N</i> 4698	(28156)	22699	27397	1232.4 ^b
		<i>T</i> 5366	(44488)	35865	41231	
		<i>p</i> 0.876		0.633		

^a Total is based on Caucasian observed (obs.) plus Japanese adjusted (adj.) value (see text)

^b $P < 0.001$

1983); Japanese plasma, $5.7\% \pm 0.7\%$ (Asakawa et al. 1985); Caucasian erythrocyte, $3.1\% \pm 0.5\%$ (Rosenblum et al. 1984); Japanese erythrocyte, $4.0\% \pm 0.3\%$ (Takahashi et al. 1986); Caucasian platelets, $2.6\% \pm 0.4\%$ (Hanash et al. 1986a); Caucasian lymphocytes, $3.1\% \pm 0.2\%$ (Hanash et al. 1986b). Thus, the present data continue the previously noted trend that in our hands, the heterozygosity indices yielded by 2-D PAGE are roughly one-half to two-thirds the indices based on 1-D E of serum proteins and erythrocyte enzymes (see Harris and Hopkinson 1972; Harris 1980). We note in passing, however, a certain arbitrariness to these indices, in that they are so technique-dependent. On the one hand, the widely quoted summary of Harris and Hopkinson (1972) does not reflect the additional variation revealed by the relatively recent application of isoelectric focusing to a number of proteins but, on the other hand, does reflect an effect to optimize for each protein the 1-D gel conditions most appropriate to the resolution of variants (whereas 2-D PAGE applies a uniform set of conditions to all polypeptides). Be all this as it may, we believe the fact that we (see also Goldman and Merrill 1983; Goldman et al. 1985) find a greater amount of genetic variation in 2-D PAGE preparations than did many of the earlier investigators primarily reflects the result of improving gel techniques, care in spot selection (see Neel et al. 1985), and the use of family material.

By now our studies with 2-D PAGE of plasma, erythrocyte lysate, and platelets have revealed a total of 13 polymorphic

polypeptides in Caucasians and 24 in Japanese, 10 of which appear to involve identical polypeptides in the two populations. The data are now sufficient to render a preliminary systematic consideration of ethnic differences of interest. Furthermore, 7 of the erythrocyte enzymes studied with 1-D E in the course of mutation rate studies in Japan (Neel et al. 1980) and Ann Arbor (Mohrenweiser et al., in press) are polymorphic in both populations; we may ask how well the "traditional" picture of ethnic variation revealed by these enzymes agrees with the findings from 2-D PAGE.

In comparing ethnic variation, a standard problem has been that the data on the various genetic systems under consideration must usually be assembled from a mosaic of populations, any one of which may or may not be representative of the remainder. We are fortunate in the present setting that all the results to be compared were obtained on the same two populations, the one Caucasians residing in Ann Arbor, the other Japanese from Hiroshima and Nagasaki. Furthermore, because so many of the techniques have been coordinated between the Ann Arbor and the Hiroshima laboratories, there was an unusual degree of comparability in the techniques in the two laboratories.

Table 2 presents the results of a simple 2×2 χ^2 comparison between the frequency of the normal allele of the 10 polypeptides thus far found to be polymorphic in 2-D PAGE preparations in both Ann Arbor and Hiroshima-Nagasaki. Because of the small numbers involved in some series, a single variant can

Table 4. A comparison of allele frequencies in Japanese and Caucasians for the four polymorphisms identified in both platelets and lymphocytes

Designation	Author	Source ^a	Ethnic ^b	Allele count			Interethnic χ^2
				N	V	Total	
C-005	This paper	P	J	160	10	170	15.44 ^c
C-38k	Kondo et al. (1985)	L	J	313	35	348	
NIMH-1	Goldman and Merrill (1983)	L	C	63	19	82	
Ly-34	Hanash et al. (1986b)	L	C	66	14	80	
C-014	This paper	P	J	156	14	170	2.35
C-31k	Kondo et al. (1984a)	L	J	391	25	416	
NIMH-16	Goldman and Merrill (1983)	L	C	64	4	68	
Ly-59	Hanash et al. (1986b)	L	C	78	2	80	
D-012 (Es-D)	This paper	P	J	103	67	170	31.03 ^c
C-33k	Kondo et al. (1984b)	L	J	115	69	184	
NIMH-10	Goldman and Merrill (1983)	L	C	50	6	56	
Ly-70	Hanash et al. (1986b)	L	C	71	9	80	
D-014	This paper	P	J	142	28	170	16.30 ^c
NIMH-6	Goldman and Merrill (1983)	L	C	44	28	72	
Ly-67	Hanash et al. (1986b)	L	C	53	27	80	

^a P, platelet; L, lymphocyte

^b J, Japanese; C, Caucasian

^c $P < 0.001$

result in an allele frequency $\geq 1.0\%$, the traditional definition of a polymorphism. For the present, we consider a system to be polymorphic only if at least two unrelated individuals with the variant have been observed. We also present in Table 3 the same type of contrast for the 7 enzymes polymorphic in both populations among the 24 enzymes studied by 1-D E by ourselves both in Caucasian children in Ann Arbor and Japanese children in Hiroshima and Nagasaki. The former sample was assembled in the course of a pilot study on the strategy of a genetic monitoring study utilizing placental cord blood samples from newborn infants; the detailed data are in press (Mohrenweiser et al., in press). The latter data set derives from a study of the potential genetic effects of the atomic bombs (Neel et al. 1980); a detailed description of the data is in preparation.

For the 2-D PAGE data four of the 10 polymorphisms present significantly different allele frequencies. For the 1-D E polymorphisms, a complication in the data is the occurrence of siblings in the Japanese study. Since, however, the sibling distribution is known, we can estimate the number of independent alleles in the sample by the procedure suggested by Chakraborty (1978). This procedure results in an "adjustment factor" that when multiplied by total alleles leads to an estimate of the number of independent alleles in the sample. Both the observed and the adjusted numbers are given in Table 3, but the χ^2 is based upon the number of independent alleles. Six among these seven contrasts present statistically significant ethnic differences. With respect to the results of both these procedures, this is of course the statistically simplest contrast possible, since it fails to take into consideration differences in allele structure in the two populations.

A major difficulty in comparing the results of 2-D PAGE with those from 1-D E is the enormous differences in sample size. In this situation, a difference of comparable magnitude will be significant in one series but not in the other. We note that the average ethnic difference between the frequency of the normal allele for the 10 polymorphisms detected by 2-D

PAGE was 0.12, whereas for the 7 studied with 1-D E, it was 0.10. This contrast suggests very little difference in the ethnic variability of the polymorphisms revealed by the two approaches.

The foregoing analysis presents only a partial picture of the ethnic differences. Thus there are 4 apparently corresponding polypeptides that present a low frequency polymorphism in one ethnic group but not the other (Japanese plasma C-001, Caucasian plasma C-001; Japanese plasma D-008, Caucasian plasma D-008; Japanese erythrocyte C-012, Caucasian erythrocyte D-009; Japanese erythrocyte D-018, Caucasian erythrocyte E-002-3). It should also be noted that 3 polypeptides found polymorphic in Caucasian populations (erythrocyte C-004, platelet PLA-1, and PLA-3) have not been found to have an identifiable counterpart in Japanese populations and, vice versa, 11 polymorphic polypeptides of the Japanese population (erythrocyte C-006, C-016, C-023, D-017, D-002 and platelet A-005, A-013, B-002, C-004, C-011, D-007) have not thus far been found to have a Caucasian counterpart. Further studies are in progress to determine the validity of these apparent differences. (The polypeptide designated C-008 in studies of plasma from Caucasoids, an apparent variant of which was suspected of presenting technical problems in scoring because of a departure from Hardy-Weinberg proportions, also appeared to behave anomalously in Japanese plasma, and was omitted from consideration, as was the D-014 and PL-3 comparison mentioned earlier.) It is not clear to what extent these findings constitute further ethnic differences, and to what extent they reflect the difficulty in being absolutely certain of spot identity in gels prepared in two different laboratories. However, from the gel position, we suspect that 3 platelet polymorphisms encountered in Japanese (A-013, C-004, and D-007), are not present in Caucasians.

There appears to be significant overlap between the platelet polymorphisms encountered in this study and certain of the recently reported lymphocyte polymorphisms. The identities in which we have confidence are shown in Table 4.

These identities are based primarily on gel position and similarities in the configuration of the surrounding spots. There are no statistically significant differences between the allele frequencies for corresponding polypeptides in the Japanese platelet and the Japanese lymphocyte series. Furthermore, the two Caucasian series appear homogeneous. [N.B. The lymphocyte samples studied by Hanash et al. (1986b) were derived from the population of Nancy, France.] When, however, we contrast the allele frequencies for the 4 platelet polymorphisms of Japanese with the allele frequencies in the corresponding lymphocyte polymorphisms in Caucasians, three of the four contrasts are significant, one of these involving the esterase D polymorphism found to be significant in the contrasts of Table 2.

By now there is sufficient experience with protein variants studied both by 1-D E and 2-D PAGE that some tentative generalizations concerning the relative efficiency of the two approaches seem appropriate. As of this writing, we are aware of studies on the resolution in a 2-D PAGE system of 34 variants, involving 8 proteins, first detected by 1-D E (Wanner et al. 1982; Fujita et al. 1985; Asakawa et al. 1984; Asakawa, unpublished). Of the 34 variants detected with 1-D E, 11 were not detected with 2-D PAGE. The reasons for failure are not only inadequate or inconstant resolution, but also "technical," as when the Hb variant "runs off" the basic end of the focusing gel. This still preliminary finding suggests that the results of 2-D PAGE should underestimate heterozygosity by almost 30%. On the assumption that in our various studies undetected variants have the same numerical representation as in the above quoted series, the appropriate correction of our own studies on plasma, erythrocyte, and platelet proteins, where the observed average index was 3.9%, would raise the index to approximately 5.1%. This fact, together with the ethnic variability described in this paper, persuades us that contrary to some earlier surmises (McConkey et al. 1979; Walton et al. 1979) the proteins revealed by 2-D PAGE differ little if at all in the amount and patterning of their genetic variation from the proteins studied by 1-D E.

These data also contribute to another consideration of some importance to population genetics. Several years ago we argued that the restriction fragment length polymorphisms (RFLPs) defined by the various nucleases were not uniformly distributed through the genome, being less common in the exon DNA than in the remainder of the genome. The argument was somewhat indirect and based in part on the frequency with which electrophoretic variants had been identified with 1-D E in serum proteins and erythrocyte enzymes (Neel 1984). Since that time Cooper et al. (1985) have strengthened the argument by demonstrating that in 19 cloned DNA segments selected at random with respect to their coding potential, the heterozygosity of single copy DNA sequence was 0.0037, approximately an order of magnitude higher than predicted from the protein data by Nei (1975). The data summarized in this paper reinforce the earlier inference concerning a relatively low frequency of RFLPs in exons, in this case presumably predominantly *not* those coding for enzymes, and strengthen the inference of much heavier selection against exon RFLPs than against the remainder of the DNA (or, less likely, lower mutation rates in the former).

Acknowledgements. We are grateful to Mrs. Y. Nagahata-Shimoichi and Miss Y. Tanaka for their technical assistance. J.V.N.'s work supported in part by Department of Energy Grant DE-AC02-82-ER-60089.

References

- Anderson NL, Hickman BJ (1979) Analytical techniques for cell fractions. XXIV. Isoelectric point standards for two-dimensional electrophoresis. *Anal Biochem* 93:312-320
- Asakawa J, Satoh C (1986) Characterization of three electrophoretic variants of human erythrocyte triosephosphate isomerase found in Japanese. *Biochem Genet* 24:131-148
- Asakawa J, Satoh C, Takahashi N, Fujita M, Kaneko J, Goriki K, Hazama R, Kageoka T (1984) Electrophoretic variants of blood proteins in Japanese. III. Triosephosphate isomerase. *Hum Genet* 68:185-188
- Asakawa J, Takahashi N, Rosenblum BB, Neel JV (1985) Two-dimensional gel studies of genetic variation in the plasma proteins of Amerindians and Japanese. *Hum Genet* 70:222-230
- Chakraborty R (1978) Appendix to Ferrell et al: Number of independent genes examined in family surveys and its effect on gene frequency estimation. *Am J Hum Genet* 30:550-552
- Cooper DN, Smith BA, Cooke HJ, Niemann S, Schmidke J (1985) An estimate of unique DNA sequence heterozygosity in the human genome. *Hum Genet* 69:201-205
- Fujita M, Satoh C, Asakawa J, Nagahata Y, Tanaka Y, Hazama R, Krasteff T (1985) Electrophoretic variants of blood proteins in Japanese. VI. Transferrin. *Jpn J Hum Genet* 30:191-200
- Goldman D, Merrill CR (1983) Human lymphocyte polymorphisms detected by quantitative two-dimensional electrophoresis. *Am J Hum Genet* 35:827-837
- Goldman D, Goldin LR, Rathnagiri P, O'Brien SJ, Egeland JA, Merrill CR (1985) Twenty seven protein polymorphisms by two-dimensional electrophoresis of serum, erythrocytes, and fibroblasts in two pedigrees. *Am J Hum Genet* 37:898-911
- Hanash SM, Neel JV, Baier LJ, Rosenblum BB, Niezgodka W, Markel D (1986a) Genetic analysis of thirty-three platelet polypeptides detected in two-dimensional polyacrylamide gels. *Am J Hum Genet* 38:352-360
- Hanash SM, Baier LJ, Welch D, Kuick R, Galteau M (1986b) Genetic variants detected among 106 lymphocyte polypeptides observed in two-dimensional gels. *Am J Hum Genet* 39:317-328
- Harris H (1980) *The principles of human biochemical genetics*, 3rd edn. Elsevier/North-Holland, Amsterdam
- Harris H, Hopkinson DA (1972) Average heterozygosity per locus in man: an estimate based on the incidence of enzyme polymorphisms. *Ann Hum Genet* 36:9-20
- Kondo I, Yamakawa K, Shibasaki M, Yamamoto T, Hamaguchi H (1984a) Genetic analysis of human lymphocyte proteins by two-dimensional gel electrophoresis. V. Genetic polymorphism of cytosol 31k polypeptide. *Hum Genet* 66:244-251
- Kondo I, Yamamoto T, Yamakawa K, Harada S, Oishi H, Nishigaki I, Hamaguchi H (1984b) Genetic analysis of human lymphocyte proteins by two-dimensional gel electrophoresis. VI. Identification of esterase D in the two-dimensional gel electrophoresis pattern of cellular proteins. *Hum Genet* 66:248-251
- Kondo I, Yamamoto T, Yamakawa K, Shibasaki M, Hamaguchi H (1985) Genetic analysis of human lymphocyte proteins by two-dimensional gel electrophoresis. VII. Genetic polymorphism of cytosol polypeptide with molecular weight of 38,000. *Hum Genet* 70:328-332
- McConkey EH, Taylor BJ, Phan D (1979) Human heterozygosity: a new estimate. *Proc Natl Acad Sci USA* 76:6500-6504
- Mohrenweiser HW, Wurzinger KH, Neel JV (in press) Frequency and distribution of rare electrophoretic mobility variants in the population of Ann Arbor, Michigan. *Ann Hum Genet*
- Morrissey JH (1981) Silver stain for proteins in polyacrylamide gels: a modified procedure with enhanced uniform sensitivity. *Anal Biochem* 117:307-310
- Neel JV (1984) A revised estimate of the amount of genetic variation in human proteins: implications for the distribution of DNA polymorphisms. *Am J Hum Genet* 36:1135-1148
- Neel JV, Satoh C, Hamilton HB, Otake M, Goriki K, Kageoka T, Fujita M, Neriishi S, Asakawa J (1980) Search for mutations affecting protein structure in children of atomic bomb survivors: preliminary report. *Proc Natl Acad Sci USA* 77:4221-4225

- Neel JV, Rosenblum BB, Sing CF, Skolnick MM, Hanash SM, Sternberg S (1984) Adapting two-dimensional gel electrophoresis to the study of human germ-line mutation rates. In: Celis JE, Bravo R (eds) *Methods and applications of two dimensional gel electrophoresis of proteins*. Academic Press, New York London, pp 259–306
- Neel JV, Baier L, Hanash SM, Erickson RP (1985) The frequency of polymorphisms for alleles encoding for proteins of the liver of domesticated mice. *J Hered* 76:314–320
- Nei M (1975) *Molecular population genetics and evolution*. North Holland, Amsterdam
- Rosenblum BB, Neel JV, Hanash SM (1983) Two-dimensional electrophoresis of plasma polypeptides reveals “high” heterozygosity indices. *Proc Natl Acad Sci USA* 80:5002–5006
- Rosenblum BB, Neel JV, Hanash SM, Joseph JL, Yew N (1984) Identification of genetic variants in erythrocyte lysate by two-dimensional gel electrophoresis. *Am J Hum Genet* 36:601–612
- Takahashi N, Neel JV, Nagahata-Shimoichi Y, Asakawa J, Tanaka Y, Satoh C (1986) Inherited electrophoretic variants detected in a Japanese population with two-dimensional gels of erythrocyte lysates. *Ann Hum Genet* 50:313–325
- Walton KE, Styer D, Greenstein EI (1979) Genetic polymorphism in normal human fibroblasts as analyzed by two-dimensional polyacrylamide gel electrophoresis. *Biol Chem* 254:7951–7960
- Wanner LA, Neel JV, Meisler MH (1982) Separation of allelic variants by two-dimensional electrophoresis. *Am J Hum Genet* 34:209–215

Received May 27, 1987 / Revised July 6, 1987