

MTR 01187

A pilot study of the use of placental cord blood samples in monitoring for mutational events *

James V. Neel, Harvey W. Mohrenweiser and Henry Gershowitz

Department of Human Genetics, University of Michigan Medical School, Ann Arbor, MI 48109 (U.S.A.)

(Received 1 December 1986)

(Accepted 2 February 1987)

Keywords: Placental cord blood; Mutant proteins; Electrophoresis, one-dimensional.

Summary

A pilot study exploring the examination of placental cord blood samples for mutant proteins with one-dimensional electrophoretic techniques is described. Although technical advances are such that the techniques employed in this study are now partially superceded, the practical problems encountered in this study would be typical of any monitoring program of this type. No mutations altering electrophoretic mobility among a battery of 51 different locus products were encountered in a total of 277 747 locus tests. When these data are combined with similar data from other studies, the mutation rate for electromorphs becomes 0.34×10^{-5} /locus/generation.

As the papers in the special number of *Mutation Research* illustrate, there now exist a number of sophisticated approaches to monitoring for genetic damage in somatic cells. Unfortunately, the value of these systems as predictive indicators for germ line damage transmitted to the next generation remains undetermined. An outstanding need in the field of human genetic monitoring is to build the necessary bridge between somatic and transmitted genetic endpoints, based insofar as possible on comparable indicators. Any really significant monitoring program of the future will undoubtedly pursue several different endpoints,

somatic and genetic, in an effort to build the bridges.

Between 1976 and 1984 we developed methods for, and engaged in, a pilot study in the U.S.A. of one such possible component of a genetic monitoring program, based on a combination of a search for charge-change variants with one-dimensional electrophoresis (1-D E) and for loss-of-activity variants by quantitative enzyme analysis. This may be considered the counterpart at the germinal level of the various efforts to develop techniques for identifying 'point mutations' in erythrocytes and lymphocytes. This study has now terminated; this presentation will undertake to summarize the data and lessons generated by the program. It is worth emphasizing at the outset that these 'lessons' will in many respects also apply to any similar study of mutation utilizing such other approaches as the examination of proteins by two-dimensional polyacrylamide gel elec-

Correspondence: Dr. James V. Neel, Department of Human Genetics, University of Michigan Medical School, Ann Arbor, MI 48109 (U.S.A.)

* Research supported by Department of Energy Contract AC-02-82-ER-60089.

trophoresis or of DNA by the various available methodologies. Preliminary references to this study have appeared in Neel et al. (1979, 1980) and Mohrenweiser (1981, 1983a).

Protocol

Justification of indicators

The spectrum of spontaneous and induced mutation at the DNA level ranges from single base pair substitutions to duplications, deletions, inversions, or translocations of large blocks of DNA, up to and including entire chromosomes. Electrophoresis as currently practiced is designed to detect that fraction of nucleotide mutations which result in the substitution in a polypeptide chain of an amino acid whose charge differs from that of its predecessor, with retention of protein function. The quantitative enzyme assays should detect mutations resulting in complete or very significant loss of enzyme activity, the causes ranging from single nucleotide substitutions resulting in stop codons or unacceptable amino acid substitutions to deletions of varying size, both involving not only exons but introns and regulatory sites (Mohrenweiser, 1983b). In principle, electrophoresis, in which the location of the protein is indicated either by a protein or an activity stain, should permit the identification of half-normal amounts of protein, but the extraneous sources of variation in the staining procedure and limitations of densitometric scanners have in the past limited the usefulness of this strategy.

Different mutagens will produce different spectra of mutation. For instance, current evidence suggests that chemical mutagens will produce relatively more base pair substitutions than will ionizing radiation while radiation may produce predominantly deletions and rearrangements. Elsewhere we are reviewing this matter in some detail (Neel et al., in preparation). While the generalization is probably correct, we would argue that in a monitoring situation it is important to attempt to capture the entire spectrum of events, so that it is as necessary to establish what has not happened as well as what has. Thus endpoints which detect a wide variety of mutational events should be incorporated into a proper monitoring effort.

Source of samples

There are of course a variety of sampling strategies for a population thought to be (or to have been) at risk of an increased mutation rate. Rarely, a population will be perceived as entering into a situation where mutation rates may be increasing; in this instance, if the population is sufficiently large, children studied for the occurrence of mutation at the beginning of the program may serve as controls for the children born and studied later. More commonly, a population will have been subjected to a defined insult (atomic bomb explosion, nuclear power plant malfunction, inadvertent exposure to mutagenic chemical), and one must develop a suitable control group to the children born of exposed parents. One could, in a study employing biochemical indicators, use children of the same parents born prior to the insult as control on the children born after the insult, but even when biochemical (including DNA) variants in the heterozygous state are the indicators, it is possible there is negative selection against some variants, so that a retrospective study may be biased. A preferable alternative is a study where controls are identified and studied contemporaneously with the children of exposed. In this design, the collection of blood samples from the placentae of newborn infants, after the cord has been severed, plus venous samples collected from both parents shortly preceding or following the delivery of the child, would appear to have much to recommend it; this was the procedure we have followed in a study of a normal population.

In industrialized countries, most deliveries are concentrated in hospitals where such collections are convenient, and the monitoring of newborns leads to the identification of any increased mutation rate at the earliest possible moment. The principal disadvantage of this particular protocol is that if one needs to resample the child for any of a number of reasons, this may be quite difficult in a home setting during the first several years of life, the more so since the justification is scientific rather than medical. Given, in addition, the mobility of the American public, and the operation of Murphy's genetic law number 12, which says that the parents of the child you most need to sample have just relocated in a distant setting, we conclude that all studies should be performed without

the expectation of second samples (although this possibility is not excluded).

The University of Michigan Medical Center has a relatively small maternity service, averaging about 1800 deliveries/year. On an annual basis, the proportion of deliveries for which it was possible to collect samples from both parents ranged from 25% to 75%. The high figure reflects the activities of a very committed Fellow in Perinatology (Dr. Linda Dunn), the low figure the lesser commitment by the house staff who followed her. An intermediate figure was obtained when the responsibility was delegated to the obstetric nursing staff, and a small fee paid into an 'educational fund' for each completed trio. This latter approach was adopted during the latter three-quarters of the study.

The reasons for not achieving a higher completion rate are various. All samples were collected with informed consent; not all couples wish to participate in a study which appears of no benefit to them or their child. The length of hospitalization following an uncomplicated delivery is now only about 3 days; it is not always possible to make contact with the father during this brief period. Finally, the collection of cord blood samples has a rather low priority whenever the delivery rooms are especially busy. The ethnic composition of the final sample of 3456 infants was: 88% Caucasians, 7% American Blacks, and 5% of Mongoloid, mixed or unknown ethnicity. Among the total sample, there were 210 sibpairs and 7 sib trios. There were also 7 apparently identical twins, of whom only one member of each pair has been included in the sample.

The decision to collect (and process) samples from both parents at the time of their child's delivery was based on the following considerations. A variant with the electrophoretic mobility of a common polymorphism has a low probability of being a newly arisen mutation. Rare (i.e., non-polymorphic) variants have a much higher relative probability of resulting from a recent mutation, and the decision was reached to focus the search for mutations on these. Our preliminary experience (Neel et al., 1978), and that of others, indicated that for any particular locus, such rare variants collectively were encountered as heterozygotes with a frequency of several per 1000 de-

terminations. Arbitrarily selecting a figure of 2 per 1000 locus determinations, we calculated that if the number of loci under study ultimately numbered 40, then the need for an examination of parental samples would arise in $100[1 - (998/1000)^{40}]$, or 7.7% of all infants studied. No comparable preliminary data were available for the frequency of heterozygotes for enzyme deficiency variants in an unselected series, but their occurrence in any frequency could only increase the need for family studies. It was 'guesstimated' that the collection and processing of parental samples close to the time of the child's birth would be more efficient in terms of time, costs, and availability, than the attempt to collect samples from the parents of some 10% of the infants only after a variant had been demonstrated in the child. Accordingly, no child's sample was examined unless specimens from both parents were also available.

A number of variants were encountered which were apparently identical in phenotype with variants known to be low-frequency polymorphisms within one of the two principal ethnic groups studied. The probability that any one of these represented a new mutation was very low. Unlike the common polymorphisms, however, which on any given working day would appear many times in the day's gels and so could be readily compared, these were encountered once or twice a week as isolated occurrences on a gel. We have accordingly elected to perform family studies on all those variants, as they were encountered (as well as the necessary comparison runs from time to time). They include the 2 type of AK, the C type of 6PGD, the C type of ACP, the D types of TF, the Michigan and A variants of CRPL, the 2 type of PEPA, and the 2 and 3 variants of PEPD. Although some would argue that family studies on these variants in a protocol investigating mutation is unnecessary, we regard it as a conservative approach to the estimation of mutation rates. These family study data will also prove useful in the context of how the issue of non-paternity enters into a study of this type and will be introduced later.

It was realized from the outset that this study was not of a magnitude to yield by itself definitive data on the frequency of mutation resulting in

electrophoretic variants. On the other hand, the data could be — and have been — pooled with other appropriate data to yield such an estimate (Neel et al., 1986). A major emphasis of the Ann Arbor effort was the development and/or adaptation of biochemical methods for detecting germinal mutation, with the satisfactory methodologies to be introduced subsequently into the laboratories of the Radiation Effects Research Foundation in Hiroshima and Nagasaki, as the basis for the much larger study of the offspring of atomic bomb survivors (Neel et al., 1979, 1980). Finally, the data resulting from the study will eventually be incorporated into a treatment of the comparative genetics of a variety of ethnic groups.

Methodology

Blood samples were collected and processed as described by Neel et al. (1979) and Fielek and Mohrenweiser (1979). An aliquot of cells was retained in a glycerol-sorbitol mixture for blood typing (Weiner and Lourie, 1981). The electrophoretic methodology has been previously described (Neel et al., 1980, 1986; Barrantes et al., 1982; Long et al., 1986; Mohrenweiser and Neel, in preparation). Quantitative enzyme assay procedures were as described by Fielek and Mohrenweiser (1979) and Mohrenweiser (1981).

Parentage verification

When a rare variant in a child was not present in either parent, the child's blood sample and those of both parents were subjected to a battery of serological tests designed to detect a discrepancy between legal and biological parentage. In addition, certain of the loci examined for the occurrence of mutation are characterized by polymorphisms useful in this context. The fact that the sample from the child was from a newborn infant precluded the use of several markers customarily used in such determinations: the subtypes of ABO, and the P, Lewis (erythrocyte), GM, KM, and HP systems. The serological systems routinely employed were ABO, MNS, RH, K, FY, JK; the biochemical typings yielded data on common polymorphisms of the ESD, GLO, ACP, AK, ADA, PGM1, GPT and GALT loci. The probability of excluding a Caucasoid male incorrectly identified as the father of a given child was 0.88

with this battery. The corresponding exclusion probability for an American Black was 0.82. In the event of a putative mutation in which there was no exclusion with this battery, further efforts at exclusion would involve isoelectric focusing for the subtypes of PGM1 and GC, and HLA typings. The latter would involve recontacting the family since aliquots of leukocytes were not routinely stored. If the family was recontacted when the infant was at least 6 months old, the tests mentioned above as impossible in a newborn could also be performed.

In addition, in order to set a prior probability of biological-legal parentage discrepancies in this sample, over a period of 3 years, 220 randomly selected trios of samples (father, mother, and child) were subjected to the above-enumerated typings. Such prior probabilities are necessary to any calculation of the relative probabilities of an apparent mutation being the result of an undetected parentage exclusion rather than a true mutation. Among this sample, there were 196 Caucasoid trios, 18 American Black trios, 3 mixed Caucasoid-Black, and 2 for whom race was not indicated.

Results

Variant frequencies

The total number of locus tests for electrophoretic variants and the frequency of rare variants is shown in Table 1. Allowance has been made for loci which are duplicated (HBA1, HBA2 and HBG1, HBG2) and proteins which are composed of two or more polypeptides (HB, LDH, ESA). A total of 51 loci are represented in the data. The varying number of determinations from locus to locus reflect the phasing in of additional systems as the study progressed. The number of variants indicated for each locus must in general be an underestimate, since no single electrophoretic system will detect all of the charge-change variants of a given protein. Some of the more interesting variants in this series have already been described [CP — Mohrenweiser and Decker (1982); TPI — Asakawa and Mohrenweiser (1982); GOT — Wurzinger and Mohrenweiser (1982); GPI — Mohrenweiser et al. (1987); G6PD — Mohrenweiser and Fielek (in preparation)]; a

TABLE 1
 FREQUENCIES OF RARE ELECTROPHORETIC VARIANTS IN A SERIES OF ELECTROPHORETIC DETERMINATIONS
 ON ERYTHROCYTE AND PLASMA PROTEINS

Enzyme	EC No.	Abbrev.	Caucasians (variants/ determs.)	Am. Blacks (variants/ determs.)	Total ^a (variants/ determs.)
Acetaldehyde dehydrogenase	1.2.1.3	ALDH	0/1256	0/99	0/1431
Acid phosphatase 1	3.1.3.2	ACPI	2/3026	4/256	7/3450
Adenosine deaminase	3.5.4.4	ADA	0/3007	0/253	0/3436
Adenylate kinase 1	2.7.4.3	AK1	4/3003	0/253	4/34320
Albumin		ALB	2/2743	0/221	2/3118
Apolipoprotein A1		APOA1	1/1076	2/76	3/1220
Carbonic anhydrase-1	4.2.1.1	CA1	0/676	2/73	2/784
Carbonic anhydrase-2	4.2.1.1	CA2	0/890	0/93	0/1031
Ceruloplasmin		CRPL	2/2938	3/246	5/3348
Diaphorase	1.6.*.*	DIA	5/696	0/47	5/790
Enolase-1	4.2.1.11	ENO1	0/1257	0/99	0/1432
Esterase A-1	3.1.1.1	ESA1	3/3020	1/255	4/3443
Esterase A-2	3.1.1.1	ESA2	1/3020	0/255	1/3443
Esterase A-3	3.1.1.1.	ESA3	0/3020	0/255	0/3443
Esterase A-C	3.1.1.1	ESAC	1/3020	0/255	1/3443
Esterase B	3.1.1.1	ESB	0/1848	0/144	0/2098
Esterase D	3.1.1.1	ESD	0/3028	0/255	0/3451
Fumarate hydratase (fumarase)	4.2.1.2	FUM	0/1071	0/78	1/1214
Gal-1-phosphate uridyl transferase	2.7.7.12	GALT	1/2969	0/249	1/3384
Glutamate-oxaloacetic transaminase(S)	2.6.1.1	GOTS	11/3024	3/255	17/3447
Glucose-6-phosphate dehydrogenase	1.1.1.49	G6PD	5/2485	0/197	6/2826
Glucosephosphate isomerase	5.3.1.9	GPI	18/3007	0/253	18/3427
Glutamate-pyruvate transaminase(S)	2.6.1.2	GPTS	6/1301	1/101	7/1481
Glyoxalase-1	4.4.1.5	GLO1	0/1402	0/126	0/1608
Hemoglobin A-1		HBA1	0/3007	3/253	3/3427
Hemoglobin A-2		HBA2	0/3007	0/253	0/3427
Hemoglobin B		HBB	1/2984	1/252	2/3403
Hemoglobin GA		HBGA	1/3007	0/253	1/3427
Hemoglobin GG		HBGG	0/3007	0/253	0/3427
Hexokinase-1	2.7.1.1	HK1	2/2938	0/248	2/3354
Hexokinase-2	2.7.1.1	HK2	0/3007	0/254	0/3429
Inosine triphosphatase	3.6.1.19	ITPA	0/1256	0/99	0/1431
Isocitrate dehydrogenase(S)	1.1.1.42	ICDS	2/2915	0/242	2/3317
Lactate dehydrogenase A	1.1.1.27	LDHA	2/3007	4/253	6/3427
Lactate dehydrogenase B	1.1.1.27	LDHB	1/3007	0/253	2/3427
Malate dehydrogenase(S)	1.1.1.37	MDHS	0/3007	0/253	0/3427
Nucleoside phosphorylase	2.4.2.1	NP	3/3019	0/256	3/3443
Peptidase A	3.4.11.*	PEPA	5/3020	0/256	5/3444
Peptidase B	3.4.11.*	PEPB	7/3025	2/257	9/3450
Peptidase C	3.4.16.*	PEPC	9/2832	1/236	10/3226
Peptidase D	3.4.13.9	PEPD	3/2834	1/238	4/3230
Phosphoglucomutase-1	2.7.5.1	PGM1	4/3028	0/257	4/3453
Phosphoglucomutase-2	2.7.5.1	PGM2	4/3027	2/257	6/3452
Phosphogluconate dehydrogenase	1.1.1.44	PGD	2/3006	0/252	2/3425

TABLE 1 (continued)

Enzyme	EC No.	Abbrev.	Caucasians (variants/ determs.)	Am. Blacks (variants/ determs.)	Total ^a (variants/ determs.)
Phosphoglycerate kinase	2.7.2.3	PGK	0/2104	0/161	0/2385
Inorganic pyrophosphatase	3.6.1.1	PP	0/684	0/72	0/793
Prealbumin		PALB	0/887	0/53	0/1004
7-S- β -Globulin		BG7S	0/693	0/45	0/783
Transferrin		TF	49/2938	0/246	49/3348
Triosephosphate isomerase	5.3.1.1	TPI	2/3005	0/253	2/3425
Uroporphyrinogen synthetase	4.3.1.8	URO	6/1989	0/192	6/2307
Totals			165/123023	30/10291	202/140259

^a Includes the miscellaneous ethnic groups as well.

comprehensive description of the electrophoretic findings, with an indication of the 'rare variant structure' of each locus, will be forthcoming (Mohrenweiser and Neel, in preparation). There were 202 rare variant phenotypes in 140 259

determinations. Since each determination involves the product of 2 genes (except for the sex-linked loci G6PD and PGK), this amounts to 202 variants in 277 747 locus tests. The difference in variant allele frequency between Caucasoids (165/123 023)

TABLE 2

FREQUENCIES OF ACTIVITY VARIANTS FOR A SERIES OF 12 ENZYMES SELECTED FOR RELATIVELY LOW COEFFICIENTS OF VARIATION

Further explanation in text.

Enzyme	EC No.	Abbrev.	Caucasians (variants/determs.)	Am. Blacks (variants/determs.)	Total ^b (variants/determs.)
Adenylate kinase	2.7.4.3	AK	2/1797	0/174	3/2078
Diphosphoglycomutase	2.7.5.3	DPGM	2/1160	0/88	3/1312
Enolase-1	4.2.1.11	ENO1	6/1166	1/88	10/1321
Glucose-6-phosphate dehydrogenase	1.1.1.49	G6PD	15/1812 (1) ^a	21/174 (10) ^a	38/2092
Glucosephosphate isomerase	5.3.1.9	GPI	5/1795	0/174	5/2076
Glutamate-oxaloacetic transaminase(S)	2.6.1.1	GOTS	2/1666	1/167	3/1935
Glyceraldehyde dehydrogenase	1.2.1.12	GAPD	0/1104	0/87	0/1255
Lactate dehydrogenase	1.1.1.27	LDHB	0/1819	0/176	0/2103
Malate dehydrogenase	1.1.1.37	MDH	1/1819	1/176	2/2103
Phosphoglycerate kinase	2.7.2.3	PGK	0/1819	0/176	0/2103
Pyruvate kinase	2.7.1.40	PK	18/1819	4/176	23/2103
Triosephosphate isomerase	5.3.1.1	TPI	9/1796	8/174	18/2077
Totals			60/19572	36/1830	105/22558

^a 10 of 21 are G6PD A⁻ and 1 of 15 is a C⁻ variant.

^b Includes the miscellaneous ethnic groups as well.

and Blacks (30/20 380) is significant ($\chi^2 = 16.10$, $df = 1$, $p < 0.001$).

The tests for activity variants are summarized in Table 2. There were 19 572 determinations distributed over 12 enzymes on samples from Caucasoid infants, or 35 513 locus tests; these yielded 60 variants. The corresponding figure for Blacks was 1830 determinations or 3310 locus tests, yielding 26 variants. The total is 95 variants in 40 921 locus tests. (The G6PD A⁻ variant in Blacks is not included in these numbers.) Again, the frequency of variants is higher in Blacks ($\chi^2 = 52.07$, $df = 1$, $p < 0.001$). This is true even if the TPI 1^o variant of Blacks (Mohrenweiser and Fielek, 1982) is excluded from the calculation.

Genetic discrepancies between child and parents

Two kinds of genetic discrepancies must be distinguished in this material, namely, those caused by 'clerical' error in the delivery room or laboratory, and those which have a biological basis. In the total series there were 23 trios in which the alleged newborn failed to exhibit fetal hemoglobin. Nine of these were clustered in such a way, and occurred over such a brief time span, that it was apparent that a new junior-level technician (who was shortly separated from the study) was simply not up to the task of sample processing. The remaining 14 were apparently due to errors of labeling in the primary collection process. Of course none of these trios is included in the series, although typings revealed that some of them could have been salvaged.

With these exclusions, among the 202 variants in 140 259 determinations there were 4 which occurred in a child, neither of whose parents exhibited the variant. In 2 of these cases there was a 2- or 3-system paternity exclusion based on the routine battery of tests described earlier. In 1 case, there was no exclusion with the routine battery, but a 2-system exclusion was obtained when the typings were extended to include isoelectric focusing of PGM1 and GC. (Both systems yielded an exclusion.) The remaining variant was of the sex-linked trait G6PD, where a mother's erythrocyte lysate repeatedly failed to show, on standard electrophoresis, a G6PD variant with reduced activity (12% of normal) observed in her son. G6PD activity was, however, reduced in the mother (52%

normal). Fresh blood samples were obtained from the mother and son, as well as the father and 2 sisters of the proband. Now the variant was observed in the mother's leukocytes, and faintly, in both the hemolysate and leukocyte cytosol of one of the daughters. We conclude that the mother is heterozygous for the variant but because of its reduced activity and the small anodal shift in the mobility of the variant isozyme, it could not be detected in the heterozygous mother. We are thus left with no putative mutations in this material. There were also no putative mutations in the series analyzed for quantitative variation.

Parentage exclusions in normative series

Two of the sample (trio) interchanges mentioned earlier were found among the 196 Caucasoid trios in which normative paternity testing was performed, and were excluded from the series. There are thus left 218 randomly selected trios subjected to paternity exclusion studies, for which there were no exclusions.

There is available from this study an additional estimate of the frequency of non-paternity in this data set. Each family study for a rare variant is in effect a test either for the occurrence of non-paternity or mutation. As noted, there were 202 such variants. In addition, there were 303 family studies with reference to the low-frequency polymorphisms, as described earlier. Among the latter, there were 4 instances in which the variant was not found in either parent and family studies resulted in paternity exclusions. Among the former, as described in the preceding section, there were 3 exclusions. The exclusion frequency is thus 7/527 or $1.33 \pm 0.50\%$.

Given the knowledge that there are paternity exclusions in this series, we can ask what frequency of paternity discrepancies can be excluded at the 5% probability level in the normative series of paternity studies, in which no exclusions were actually observed. With a one-tailed probability test, this is 1.4%. There is thus no conflict between the results of the normative series for paternity exclusions and the frequency observed in the family studies of the 525 electrophoretic mobility variants just described. There seems no reason why the two series cannot be combined to yield a best estimate of the frequency of extramaritally

conceived children in this study; the resulting frequency is 7/748, or $0.94 \pm 0.35\%$. Correction for the fact that our routine exclusion procedures were only 88.3% effective brings that figure to $1.11 \pm 0.38\%$.

Discussion

We will first discuss some of the facts generated by this study, and then turn to a consideration of the future of this approach as a component in an actual monitoring situation, most likely involving a study of two samples of newborn infants or older children, one set born to parents thought to be at mutagenic risk, the other to a properly matched set of non-exposed parents.

The average allele frequency of rare variants in this series was 0.8 per 1000 locus tests. This figure cannot be directly compared with any other because of differences in the loci studied in the various series. The most nearly comparable data come from the studies on the genetic effects of the atomic bombs in Japan, where a rare variant frequency of 0.8 per 1000 locus tests was also obtained among the children of parents not exposed to the radiation of the atomic bombs (Neel et al., 1986). What is important in the present context, for calculating work load, is that for any given locus, about 1 in 500 determinations will result in a finding that must be followed by family studies. With 50 loci under scrutiny, this implies that about 1 in 10 children would exhibit a finding requiring family studies. If one also did family studies on the least common polymorphisms for the reasons stated above, about 1 in 5 children would require a family study.

As noted earlier, family studies were not performed on the common polymorphisms. This convention carries with it the risk of missing a mutation whose phenotype mimics a common polymorphism. Elsewhere we have calculated that the underestimation of mutation rates which results from this omission should be about 10% (Neel et al., 1983). On the other hand, undertaking a family investigation to test whether a common polymorphism was the result of a mutational event in a parent would have increased the work load of family studies by an order of magnitude based on the actual polymorphism frequencies in these data.

A known 10% error is deemed acceptable under these conditions.

The frequency of paternity exclusions in this material appears unusually low. We have conducted two other studies which yield data on the frequency of extramarital conceptions among children born in the state of Michigan. In an analysis of a series of nuclear families from the lower portion of the state of Michigan, studied between 1956 and 1964 because of the occurrence of a potentially lethal disease in a child (Gershowitz and Neel, 1965), Schacht and Gershowitz (1963) calculated that $1.48 \pm 0.32\%$ of 1417 Caucasoid children were conceived extramaritally, the corresponding figure for 523 Black children being $9.94 \pm 1.31\%$. In a second study which involved a large series of typings of blood samples drawn in 1965 from Caucasoid individuals residing in a Michigan town with approximately 10 000 inhabitants, we estimated that $3.82 \pm 0.34\%$ of the children for whom there were typings on the nuclear family were conceived extramaritally (Sing et al., 1971). Both of these studies included the total available sibship. The frequency of detected extramarital conceptions was significantly higher among first-born children, suggesting the possibility that these exclusion figures are biased upwards by children actually conceived before marriage. The circumstances of the present study would imply that the children from whom samples were obtained were predominantly conceived in wedlock; further, in the relatively sophisticated town of Ann Arbor, parents might be reluctant to give the necessary informed consent if there were doubts concerning paternity. The earlier figures may also be inflated to some small extent by typing errors, since in both these studies 1-system exclusions (always based on duplicate typings) were accepted, and some laboratory error cannot be excluded. We thus are unwilling to engage in statistical contrasts between the various series. Let us assume, that in the U.S.A., the frequency of extramarital conceptions is between 0.01 and 0.02. As we have seen, the probability such a conceptus would receive a rare variant from the father is approximately 0.001. Finally, with a combination of the appropriate blood typings and HLA determinations, the probability of non-detection of an extramaritally conceived child is approximately

0.01 (Neel et al., 1986). Thus the probability that an apparent mutation is actually due to an undetected extramarital conception in our setting is between $0.01 \times 0.001 \times 0.01$, or 1×10^{-7} , and $0.02 \times 0.001 \times 0.01$, or 2×10^{-7} .

There were no mutations in 276 747 locus tests for electromorphs. Among the 3 studies which may be considered comparable to this one, Harris et al. (1974) reported no mutations in 133 478 locus tests on residents of the United Kingdom, Altland et al. (1982a, b) encountered 1 mutation among $\approx 225\,000$ tests on newborns in West Germany, and we encountered 3 mutations in 539 170 locus tests on children born in Hiroshima and Nagasaki to parents who did *not* receive excess radiation at the time of the atomic bombing [i.e., the 'controls' of a study on the genetic effects of the bombs (Neel et al., 1986)]. Our present findings are consistent with the results of these other studies; with this paper, the current estimate of the spontaneous rate with which mutation gives rise to electromorphs becomes 0.34×10^{-5} /locus/generation. Since only one-third to one half of nucleotide substitutions in exons result in an electrophoretic variant, these findings correspond to a nucleotide rate per (exon) locus per generation of approximately 1×10^{-5} , and per nucleotide per generation, of approximately 1×10^{-8} . (This latter calculation assumes that the average polypeptide included in this study requires 1000 exon nucleotides for its specification.) This estimate does not include gene-inactivation mutations or small duplications and deletions. The calculation of the previous paragraph suggests that false mutations resulting from undetected extramarital conceptions should make a relatively small contribution to this estimate. Inasmuch as no spontaneous mutations resulting in loss of enzyme activity have as yet been encountered, a rate cannot be calculated for this phenomenon; from experimental studies we surmise that such mutations will at least equal in frequency the above-mentioned *corrected* rate for nucleotide substitutions.

It is apparent that a search for an increase in mutation rates based on electromorphs and loss-of-functional-product mutations will involve large numbers of observations and be quite laborious, the size of any study of course depending on the desired precision. A guiding consideration in the

future for a monitoring program must be to obtain more data per individual. This is because in any actual study of the offspring of potentially mutagenized individuals, the 'front end' costs involved in the initial contacts and obtaining the blood samples are apt to make a very significant contribution to the total cost of the study, and more importantly the number of offspring born to mutagenized parents will — it is to be hoped — usually be so small that as much as possible must be learned from each offspring. In this final section, we would like to consider briefly the prospects for obtaining more data with the approaches presented in this paper and then contrast these prospects with those inherent in two-dimensional polyacrylamide gel electrophoresis (2-D PAGE), plus those arising from the first glimmerings of a DNA approach.

With respect to improvements in the 1-D E technologies employed in this study, the most significant advances would include the increased utility of isoelectric focusing, the increased availability of antibodies to a large number of proteins, and improved instrumentation for densitometric analysis. The increased utility of isoelectric focusing stems from the availability of precast gels, the ability to construct gels with gradients designed for maximum sensitivity, and automation of several steps (Altland et al., 1982a, b; Chrambeck et al., 1980; Vogel and Altland, 1982). For a study to detect mutations which alter the electrophoretic mobility of proteins, isoelectric focusing has the advantage, for many proteins, of being the most sensitive separation technique. Therefore, the minimum number of amino acid interchanges will be missed. It also has significant potential for automation. Isoelectric focusing will not totally replace other electrophoretic methods using polyacrylamide, starch, cellulose acetate, etc., as some proteins are denatured or inactivated at their isoelectric point.

Historically, most proteins have been identified, following electrophoresis, by standard protein stains if they were among the prevalent proteins in the sample or by a functional assay, i.e., enzymatic activity. The availability of antibodies has increased the number of proteins which can be studied by electrophoresis (Chapuis-Cellier et al., 1980; Tsang et al., 1983). The utilization of im-

munochemical techniques could potentially double the number of gene products which could be studied in routine blood samples. For any future program involving 1-D E, it would probably be advisable to employ the techniques of Altland to the extent possible, supplemented by those employed in this study when necessary.

New instruments devoted to densitometry, such as the LKB XL Laser Densitometer, have improved the technical feasibility of quantifying the staining intensity of proteins previously separated by electrophoresis. The assay could be a functional assay, as for enzymes, or could be linked to the immunological detection methods. This would have an advantage, in that proteins with similar enzymatic activities, but products of different loci (e.g. PGM 1 and 2), could now be quantified. It will be necessary for the sample application steps to be more accurately controlled if the proteins on the gels are to be analyzed quantitatively, although, as several enzymes/proteins are stained on each gel, it will be possible to compare the staining intensity of one enzyme to several others as references in the search for 50% staining intensity.

As regards the quantitative enzyme assays whose results are described in this paper, it should now be possible to increase the battery from the 12 we have employed to at least 25–30 erythrocyte enzymes. A series of recent developments in quantitative enzyme analyses (e.g., better computer interfacing and refined sample allocation procedures) would increase the rate of data generation using the quantitative techniques employed in generating the data described in this paper. Thus, the combination of new methods and new instruments would increase the data base per sample and reduce the cost per determination. The constraint of limited biological variation among individuals is important for quantitative analysis irrespective of the analytical technique employed. Therefore quantitative analysis will generally be restricted to the proteins of blood cells rather than plasma as quantitative changes in plasma proteins generally reflect physiological and/or pathological conditions rather than genetic regulation.

Turning now to the potential applications of 2-D PAGE to monitoring, we have recently laid out in some detail not only the current laboratory

procedures (Neel et al., 1984) but the current state of computer algorithms designed to read these gels with minimal operator intervention (Skolnick and Neel, 1986). The ability of this technique to detect electrophoretic variants has been adequately demonstrated (summary in Takahashi et al., 1987). For reasons which remain debatable, the amount of genetic variation encountered in these gels is somewhat less than with 1-D E, a fact which must be borne in mind in planning any program to study mutation rates. The principal deficiency in the technique at present is the failure thus far, in a screening mode (single densitometric readings of single gels with light optics), to develop a battery of spots in which one can distinguish with the necessary accuracy stained polypeptides with only 50% of the normal amount of protein (i.e., 'null' heterozygotes). There is, however, promise that this problem can be met with the development of thin-layer gels (Görg et al., 1985) and laser densitometry, at least for a subset of the quantitatively less variable proteins. The principal drawback with laser densitometry as currently available is the slow reading time.

With respect to cost, our current best estimate is that the cost, per research-quality 2-D gel scored, is \$20.00. In a monitoring system depending on blood samples, gels can be prepared from plasma, erythrocytes, platelets, lymphocytes, and polymorphonuclear leukocytes. There is an as yet undetermined overlap in the polypeptides visualized in the latter 3 preparations. Let us assume that there are an average of 100 different proteins which can be scored with the requisite accuracy in preparations of each of the first 3 above-mentioned blood components. Then the cost per determination is \$0.20. However, with allowance for the 'front end' costs, the costs per determination would be more like \$0.30. This cost does not include the data management system, the expense of which will vary with the size of the project. An important potential advantage of 2-D PAGE over 1-D E is the possibility that both qualitative and quantitative variation for a large number of proteins can be studied in the same 2-D gel. Our present practice in the study of genetic variation in 2-D gels is to make preparations on the mother and father of each subject simultaneously with the preparation on the child (cf. Neel et al., 1984). If this were

continued indefinitely, it would triple the number of gels necessary to a monitoring program. It is, however, our expectation to develop a standard 2-D map which includes the positions of polymorphic variants, in which case preparations from the parents would be necessary only when a rare variant was encountered, that is in approximately 1 of 4 families.

It is of some interest to attempt to cost out an appropriately sized study. Both the estimates of cost per determination for 1-D E and 2-D PAGE will decrease to some extent in the future, with automation and other technical improvements, but we must work with present costs. Let us assume a situation in which only 1 parent had been exposed to a mutagen. If, for instance, one required a minimum of 2 000 000 locus tests in the children of exposed and the same number for controls (the latter requiring, since both loci are relevant, half as many persons), the cost of a 2-D PAGE study is roughly \$900 000 on the assumption that routine gels on the parents will become unnecessary in the future. But since these costs do not include the various legitimate types of incidental overhead expenses, or the data-handling costs, or the special biochemical studies which might be undertaken to verify a mutation, the cost should at least be doubled to approximately 2 million.

The cost for the combined 1-D electrophoretic and quantitative enzyme assays in the Ann Arbor study is \$0.75 per locus test or \$1.50 per determination. Thus, to complete a study as described above using 1-D electrophoresis and quantitative assays would cost approximately 3 million dollars. As one goal of this study was to develop methods which would contribute to the program at the Radiation Effects Research Foundation, that is, a monitoring program, maximizing the data base per proband was an important consideration when the effort was initiated. The estimated cost per test completed in this study includes not only laboratory personnel and reagent costs but also includes \$15 as the cost of obtaining the sample and the cost, including personnel, for data storage and management. As noted, utilizing the methodology developed by Altland et al. (1982a, b) where feasible, could substantially decrease costs but by an amount

which cannot be estimated. Modifying the methods, for example, to incorporate more isoelectric focusing techniques and adding new loci studied by immunochemical detection methods, would not have a significant impact on cost per test but would increase the data obtained for each proband as incorporation of the automation for electrophoresis or utilization of the new generation of quantitative enzyme analyzers could reduce the personnel cost as the screening becomes more efficient.

In any future study we would suggest serious consideration be given to the routine preservation of some or all of the leukocyte fractions, for reasons ranging from the possibility of additional genetic tests to establishing cell lines for special studies when indicated. This would of course substantially increase the initial costs of the study.

There are first glimmerings of DNA systems for the study of mutation (cf. Delehanty et al., 1986). These technologies are so early, and advancing so rapidly, that it is premature to attempt to derive even soft cost figures. However, it may be pointed out that a single, average 2-D gel, especially if accurate spot quantification becomes available, would reflect the results of certain types of mutation at 200 000 nucleotides (100 spots \times 2 alleles \times 1000 exon nucleotides), and so can provide a standard against which to measure the efficiency of a DNA system.

In concluding, we would like to emphasize that even with the technical developments which may confidently be anticipated, the study of germinal mutation, whether by the techniques mentioned in this paper or by cytogenetic or clinical approaches, will remain laborious and demanding for the immediate future. There are no potentially dramatic short-cuts in sight. But without these studies, the predictive value of somatic cell studies will forever remain obscure. Once the necessary 'bridges' have been built between the somatic and genetic endpoints, one may no longer need the very demanding germ-line studies, but without these bridges, the predictive value of somatic cell studies for genetic damage for future generations remains at best dubious but at worst, a source of free-floating anxiety to already concerned people who deserve better.

Acknowledgement

These studies were supported by U.S. Department of Energy Contract AC-02.82-ER-60089.

References

- Altland, K., M. Kaempfer and M. Forssbohm (1982a) Mass screening technique for detecting globin variants from newborn dried blood samples, *Progr. Mutation. Res.*, 3, 153-157.
- Altland, K., M. Kaempfer, M. Forssbohm and W. Werner (1982b) Monitoring for changing mutation rates using blood samples submitted for PKU screening, in: B. Bonne-Tamir (Ed.), *Human Genetics, Part A: The Unfolding Genome*, Liss, New York, pp. 276-287.
- Asakawa, J., and H.W. Mohrenweiser (1982) Characterization of two new electrophoretic variants of human triosephosphate isomerase: Stability kinetic and immunological properties, *Biochem. Genet.*, 20, 59-76.
- Barrantes, R., P.E. Smouse, J.V. Neel, H.W. Mohrenweiser and H. Gershowitz (1982) Migration and genetic infrastructure of the Central American Guaymi and their affinities with other tribal groups, *Am. J. Phys. Anthropol.*, 58, 201-214.
- Chapuis-Cellier, C., A. Francina and P. Arnaud (1980) Immunofixation after electrofocusing: Application to the study of microheterogeneity and polymorphism of serum proteins, in: B.J. Radola (Ed.), *Electrophoresis '79*, de Gruyter, Berlin, pp. 711-725.
- Chrambeck, A., L. Hjilmeland and N.Y. Nguyen (1980) Gel electrofocusing with increased degrees of freedom, in: B.J. Radola (Ed.), *Electrophoresis '79*, de Gruyter, Berlin, pp. 3-22.
- Delehanty, J., R.L. White and M.L. Mendelsohn (1986) Approaches to determining mutation rates in human DNA, *ICPEMC Meeting Report No. 2, Mutation Res.*, 167, 215-232.
- Fielek, S., and H.W. Mohrenweiser (1979) Erythrocyte enzyme deficiencies assessed with a miniature centrifugal analyzer, *Clin. Chem.*, 25, 384-388.
- Gershowitz, H., and J.V. Neel (1965) The blood groups and secretor types in five potentially fatal diseases of Caucasian children, *Acta Genet.*, 15, 261-308.
- Görg, A., W. Postel, S. Gunther and J. Weser (1985) Improved horizontal electrophoresis with hybrid isoelectric focusing in immobilized pH gradients in the first dimension and laying-on transfer to the second dimension, *Electrophoresis*, 6, 599-604.
- Harris, H., D.A. Hopkinson and E.B. Robson (1974) The incidence of rare alleles determining electrophoretic variants: Data on 43 enzyme loci in man, *Ann. Hum. Genet.*, 37, 237-253.
- Long, J.C., J.M. Naidu, H.W. Mohrenweiser, H. Gershowitz, P.L. Johnson, J.W. Wood and P.E. Smouse (1986) Genetic characterization of Gainj- and Kalam-speaking peoples of Papua New Guinea, *Am. J. Phys. Anthropol.*, 70, 75-96.
- Mohrenweiser, H.W. (1981) Frequency of enzyme deficiency variants in erythrocytes of newborn infants, *Proc. Natl. Acad. Sci. (U.S.A.)* 78, 5046-5050.
- Mohrenweiser, H.W. (1983a) Biochemical approaches to monitoring human populations for germinal mutation rates, II. Enzyme deficiency variants as a component of the estimated genetic risk, in: F.J. de Serres and W. Sheridan (Eds.), *Utilization of Mammalian Specific Locus Studies in Hazard Evaluation and Estimation of Genetic Risk*, Plenum, New York, pp. 55-69.
- Mohrenweiser, H.W. (1983b) Enzyme-deficiency variants: Frequency and potential significance in human populations, *Isozymes: Curr. Topics Biol. Med. Res.*, 10, 51-86.
- Mohrenweiser, H.W., and R.S. Decker (1982) Identification of several electrophoretic variants of human ceruloplasmin including CP* Michigan, a new polymorphism, *Hum. Hered.*, 32, 369-373.
- Mohrenweiser, H.W., and S. Fielek (1982) Elevated frequency of carriers for triosephosphate isomerase deficiency in newborn infants, *Ped. Res.*, 16, 960-963.
- Mohrenweiser, H.W., and S. Fielek, Identification and characterization of a series of rare G6PD variants in a human newborn series, in preparation.
- Mohrenweiser, H.W., and J.V. Neel, Distribution of rare electrophoretic mobility variants in a human population, in preparation.
- Mohrenweiser, H.W., P.T. Wade and K.H. Wurzinger (1987) Characterization of a series of electrophoretic and enzyme activity variants of human glucosephosphate isomerase, *Hum. Genet.*, in press.
- Neel, J.V. N. Ueda, C. Satoh, R.E. Ferrell, R.J. Tanis and H.B. Hamilton (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., H. Mohrenweiser, C. Satoh and H.B. Hamilton (1979) A consideration of two biochemical approaches to monitoring human populations for a change in germ cell mutation rates, in: K. Berg (Ed.), *Genetic Damage in Man Caused by Environmental Agents*, Academic Press, New York, pp. 29-47.
- Neel, J.V., H.W. Mohrenweiser and M.M. Meisler (1980) Rate of spontaneous mutation at human loci encoding protein structure, *Proc. Natl. Acad. Sci. (U.S.A.)*, 77, 6037-6041.
- Neel, J.V., H. Mohrenweiser, S. Hanash, B. Rosenblum, S. Sternberg, H. Wurzinger, E. Rothman, C. Satoh, K. Goriki, T. Krasteff, M. Long, M. Skolnick and R. Krzesicki (1983) Biochemical approaches to monitoring human populations for germinal mutation rates, I. Electrophoresis, in: W. Sheridan and F.J. de Serres (Eds.), *Utilization of Mammalian Specific Locus Studies in Hazard Evaluation and Estimation of Genetic Risk*, Plenum, New York, pp. 71-93.
- Neel, J.V., B.B. Rosenblum, C.F. Sing, M.M. Skolnick, S.M. Hanash and S. Sternberg (1984) Adapting two-dimensional gel electrophoresis to the study of human germ-line mutation rates, in: J.E. Celis and R. Bravo (Eds.), *Two-Dimensional Gel Electrophoresis of Proteins*, Academic Press, New York, pp. 259-306.
- Neel, J.V., C. Satoh, K. Goriki, M. Fujita, N. Takahashi, J. Asakawa and R. Hazama (1986) The rate with which

- spontaneous mutation alters the electrophoretic mobility of polypeptides, *Proc. Natl. Acad. Sci. (U.S.A.)*, 83, 389–393.
- Neel, J.V., H.W. Mohrenweiser, E.D. Rothman and J.M. Naidu (1986) A revised estimate of mutation rates in Amerindians, *Am. J. Hum. Genet.*, 38, 649–666.
- Neel, J.V., C. Satoh, N. Takahashi, J. Asakawa, M. Fujita, and R. Hazama, Search for mutations affecting protein structure and/or activity in children of atomic bomb survivors: Final report, in preparation.
- Schacht, L.E., and H. Gershowitz (1963) Frequency of extramarital children as determined by blood groups, in *Proc. Second Int. Congr. of Human Genetics, Istituto 'Gregorio Mendel', Rome*, pp. 894–897.
- Sing, C.F., D.C. Shreffler, J.V. Neel and J.A. Napier (1971) Studies on genetic selection in a completely ascertained Caucasian population, II. Family analyses of 11 blood group systems, *Am. J. Hum. Genet.*, 23, 164–198.
- Skolnick, M.M., and J.V. Neel (1986) An algorithm for comparing two-dimensional electrophoretic gels, with particular reference to the study of mutation, *Adv. Hum. Genet.*, 15, 55–160.
- Takahashi, N., J.V. Neel, Y. Nagahata-Shimoichi, J. Asakawa, Y. Tanaka and C. Satoh (1987) Inherited electrophoretic variants detected in a Japanese population with two-dimensional gels of erythrocyte lysates, *Ann. Hum. Genet.*, 50, 313–325.
- Tsang, V.C.W., J.M. Peratta and A.R. Simmons (1983) Enzyme-linked immunoelectrotransfer blot techniques for studying the specificities of antigens and antibodies separated by gel electrophoresis, *Methods Enzymol.*, 92, 377–391.
- Vogel, F., and K. Altland (1982) Utilization of material from PKU-screening programs for mutation screening, *Progr. Mutation Res.*, 3, 143–152.
- Weiner, J.S., and J.A. Lourie (1981) *Practical Human Biology*, Academic Press, New York.
- Wurzinger, K.H. and H.W. Mohrenweiser (1982) Studies of the genetic and nongenetic (physiological) variation of human erythrocyte glutamic oxaloacetic transaminase, *Ann. Hum. Genet.*, 46, 191–201.