Jpn. J. Genet. (1985) **60**, *pp.* 167-198

The estimate of protein polymorphism in human populations: lack of evidence for overestimation due to post-translational modification

By Paul A. FUERST* and Robert E. FERRELL**

* *Department of Genetics, The Ohio State University Columbus, Ohio* 43210 *USA*
** *Department of Human Genetics, University of Pittsburg School of Public Health Pittsburg, Pennsylvania USA*

(*Received December* 1, 1983)

ABSTRACT

New sensitive population screening procedures, such as isoelectric focusing, enzyme activity assays and the use of multiple electrophoretic conditions have uncovered considerable hidden variation at many structural loci. It has been suggested, however, that a significant portion of the newly identified variation is due to alleles at modifier genes which cause the post-translational modification of the structural gene product. Some investigators have questioned whether genetic variability assigned to structural loci in *Drosophila* has been greatly overestimated because of a failure to recognize the existence of such polymorphic modifier genes. Data from human populations has been reviewed for evidence of such genetically mediated posttranslational modification (*GMPTM*). Tnree genetic systems (glucose-6-phosphate dehydrogenase, alpha-1-antitrypsin and hemoglobin) are reviewed in detail. Each system is subject to non-genetic post-translational modification, but none of the data currently available indicates any role of *GMPTM*. Other data, including studies on null alleles, activity variants, population surveys and somatic cell hybrids is also reviewed. Again, no convincing evidence has been uncovered for the presence of *GMPTM*. It is concluded that the current estimates of polymorphism for enzymes and structural proteins in human populations have not been inflated due to the presence of polymorphism for modifier loci.

1. INTRODUCTION

Until very recently, the study of biochemical genetic variation in human populations has relied predominantly upon electrophoretic screening techniques to detect allelic variation. Average heterozygosities in different major racial groups have been estimated to lie in the range 6-10% (Harris and Hopkinson 1972; Nei and Roychoudhury 1974, 1982). Lately, however, the application of electrophoresis for the screening of genetic variation in natural populations has dramatically changed. New techniques have been introduced which promise to detect more of the "hidden" variability which is not detected by standard methods of starch gel or polyacrylamide gel electrophoresis.

These innovations include isoelectric focusing in polyacrylamide gels, various sequential screening procedures, two-dimensional electrophoresis, and non-electrophoretic techniques which depend upon biochemical properties of allelic variants such as heat stability or substrate specificity. Some of these methods may increase the estimated average heterozygosity in a population, compared to the values obtained by more traditional screening methods, because they separate previously undifferentiated alleles (Ayala 1982), while others may have an unknown effect because they are increasing the number and type of loci which are being sampled, but with an undetermined sensitivity to allelic differences (McConkey, Taylor and Phan 1979; McLellan, Ames and Nikaido 1983). Nei and Roychoudhury (1982) reanalyzed the data on genetic variation in major racial groups and reported that average heterozygosity levels for protein loci were increased to the range 12–16% when they included loci for which new screening methods (excluding two-dimensional electrophoresis) had been used during the collection of gene frequency data.

Some investigators question, however, whether all the variability being uncovered by new screening methods is really due to the structural locus which we think is being studied (Johnson 1979). Finnerty and Johnson (1979) and Finnerty, McCarron and Johnson (1979), presented results gathered by the method of gel sieving (Johnson 1975) which strongly suggested that genetically determined post-translational modification contributes to the levels of electrophoretic variability previously attributed to the structural loci for xanthine dehydrogenase and aldehyde oxidase in *Drosophila melanogaster*. This post-translational modification is mediated by genes other than the structural loci under consideration. Furthermore, these workers suggested that such systems may be responsible for a presently undetermined, but significant portion of the observed electrophoretic polymorphism in natural poputions which we now ascribe to variation at many structural gene loci. Specifically, they state that "much of the enzyme variation reported in natural populations may reflect modifications of primary gene products mediated by one or more other loci" (Finnerty and Johnson 1979, p. 696). Further studies of other loci by this same group suggest that polymorphism attributable to post-translational modification may occur at most loci in *Drosophila melanogaster* (Johnson 1979).

The experimental assessment of levels of genetic variation in natural populations, and attempts to understand the physiological, ecological and evolutionary meaning of this variation continue to be central goals in population genetics. Demonstration that a significant fraction of the reported variation at structural gene loci is actually due to variation at modifier loci would have profound implications for both experimental and theoretical population genetics. Data have recently been presented that genetically mediated post-translational modification does exist in natural populatiosn of

Drosophila and E. coli (Johnson *et al.* 1980; Johnson, Finnerty and Hartl 1981; Dykhuisen *et al.* 1985).

Harris (1969), Hopkinson and Harris (1971), Turner, Fisher and Harris (1975) and Harris and Hopkinson (1977) have amply documented that post-translational modification is an important factor determining the isozyme patterns observed for human enzymes. They reviewed several mechanisms which cause the production of secondary isozymes, but failed to report any examples which precisely coincide with the situation described in *D. melanogaster*. Post-translational modification certainly is occuring and does seem to play a central role in normal cell aging (Dreyfus, Kahn and Schapira 1978). In this paper we will review data on genetic variation for several loci in human population. We believe that these data are informative for judging the extent to which modifier loci contribute to current estimates of the levels of structural genetic variation. In this paper we will be concerned only with the potential existence of post-translational modifications which affect variation for enzyme or structural protein loci; we have not condsidered here any post-translational alteration of immunological or blood group variability, for which the structural basis of variation may be less clear.

2. ASSUMPTIONS OF THE ANALYSIS

Johnson (1979) and Finnerty and Johnson (1979) listed several observations which led them to propose that genetically mediated post-translational modification, abbreviated as *GMPTM*, is making an important contribution to the levels of measured electrophoretic variability in natural populations. Their interpretations are most relevant to variability which has not normally been detectable by using a single "standard" screening method, and which has therefore been classified until recently as "hidden" variation. In human surveys, this class would include variation which is detected by multiple electrophoretic or isoelectric focusing techniques, by examination of kinetic or stability properties of the protein, or by the structural analysis (amino acid or nucleic acid sequence) of a variant. Johnson (1979) and Finnerty and Johnson (1979) identified three "unresolved" problems which affect our interpretation of data at many loci: (1) levels of genic variability which are higher than predicted by any of the current theories of population genetics, (2) deficiency of detectable heterozygotes for many newly discovered variants, and (3) the existence of "too many" rare alleles. We will address each of these problems. Further implications of *GMPTM* were raised in a series of letters (Coyne, Eanes and Lewontin 1979; Singh 1979; Johnson and Finnerty 1979a,b). Discussion focused on the Mendelian segregation patterns which are expected to occur when variability exists in the population at both structural and regulatory loci, and the related problem of genetic interpretation when

modification is found to be dependent upon the presence of a specific structural allele, but does not occur for other alleles ("allele specificity of modifier").

Finnerty and Johnson (1979) presented a detailed discussion of the isozyme patterns expected when variability exists at the structural locus or at the modifier locus, but did not deal with the situation in which variability occurs at both. In view of the structural variability which is estimated to exist at a large proportion of the enzyme loci in man, it is important to consider the isozyme patterns expected when heterozygosity exists within a family for both structural and regulatory loci. First, following the arguments of Finnerty and Johnson (1979) and Johnson (1979), we assume that modification will occur in a dominantly inherited fashion. It is difficult to envision a codominant modifier, unless such a modifier was closely linked to the structural locus and affected only the cis-linked structural allele (Coyne, Eanes and Lewontin 1979). We also assume that variability at the modifier locus is sufficiently rare that we can usually ignore the presence of homozygosity for modifiers. Finally, we assume, following Coyne, Eanes and Lewontin (1979), that the modifier affects all structural alleles at a locus, and changes the mobility of any particular structural allele in such a way that a "new" variant electromorph would be produced and attributed to the structural locus. Allele specificity of modifiers (Johnson and Finnerty 1979a) might occur, but unless *most* common alleles are subject to modification it would be difficult to conclude that *GMPTM* is very important. Under the null hypothesis of no *GMPTM*, electrophoretic variability would normally be inherited in a codominant fashion (Harris and Hopkinson 1977).

To investigate the inheritance of potential modifier variability, we must specify the manner in which a modifier will alter the phenotypic ratios which appear in a pedigree. The expected progeny phenotypes for various matings involving structural and/or modifier heterozygosity are given in Tables 1 and 2 for the case of a sex-linked and an autosomal structural locus, respectively. The case of an X-linked modifier locus can easily be worked out and has not been presented here. The majority of data from humans which bears on the question of *GMPTM* does not usually contain extensive family pedigree data, but usually consists of information about sets of single families. Consequently, the phenotypic expectations in Tables 1 and 2 are not presented in terms of progeny frequencies, but rather simply in terms of the potential phenotypes which could result from any particular mating.

It can be seen in both tables that all matings which involve heterozygosity for the modifier locus will produce some "new" progeny phenotypes which are inconsistent with an hypothesized codominant transmission of allelic variation for the structural locus. Even if we assume that there is dominantly inherited structural variation, all matings in which both loci are variable should produce unexpected progeny (when compared to structural variability

Table 1. *Potential matings and expected offspring phenotypes when an X-linked structural gene product is subjected to genetically mediated post-translational modification*

	Male parent		Female parent		Progeny phenotypes	
	Genotype	Phenotype	Genotype	Phenotype	Male phenotypes	Female phenotypes
	Genetic variation at structural locus; no genetic variation at modifier locus					
1.	*N/Y ; m/m*	N	*N/N ; m/m*	N	N	N
2.	*N/Y ; m/m*	N	*N/V ; m/m*	N/V	N, V	N, N/V
3.	*V/Y ; m/m*	V	*N/N ; m/m*	N	N	N/V
4.	*V/Y ; m/m*	V	*N/V ; m/m*	N/V	N, V	N/V, V
	No genetic variation at structural locus; genetic variation at modifier locus					
5.	*N/Y ; m/m*	N	*N/N ; M/m*	(NM)	N*, (NM)	N**, (NM)**
6.	*N/Y ; M/m*	(NM)	*N/N ; m/m*	N	N, (NM)*	N**, (NM)**
7.	*N/Y ; M/m*	(NM)	*N/N ; M/m*	(VM)	N***, (NM)	N***, (NM)
	Genetic variation at both structural and modifier loci					
8.	*N/Y ; M/m*	(NM)	*N/V ; m/m*	N/V	N, V, (NM)* (VM)***	N****, N/V****, (NM)**, (NM)/(VM)***
9.	*N/Y ; m/m*	N	*N/V ; M/m*	(NM)/(VM)	N*, V***, (NM), (VM)	N**, N/V***, (NM)**, (NM)/VM)****
10.	*V/Y ; m/m*	V	*N/N ; M/m*	(NM)	N***, (NM)	N/V***, (NM)/(VM)***
11.	*V/Y ; M/m*	(VM)	*N/N ; m/m*	N	N, (NM)***	N/V***, (NM)/(VM)***
12.	*V/Y ; M/m*	(VM)	*N/V ; M/m*	(NM)/(VM)	N***, V***, (NM), (VM)	N/V***, V***, (NM)/(VM), (VM)

N = "normal" structural allele (co-dominant)
V = "variant" structural allele (co-dominant)
m = "non-modifier" allele at regulatory locus (recessive)
M = "modifier" allele at regulatory locus (dominant)
(NM) or (VM) = phenotypes resulting from modification of normal or variant structural gene products.
Genotypes given in italics; phenotypes given in non-italics

*—apparent male-to-male transmission; too many phenotypes in male progeny
**—absence of heterozygous progeny phenotypes for the expected structural phenotype; presence of unexpected homozygous progeny
***—appearance of progeny phenotype not present in either parent
****—apparent failure of Mendelian transmission from the male parent

Table 2. *Potential matings and expected offspring phenotypes when an autosomal structural gene product is subjected to genetically mediated post-translational modification*

	Parent No. 1		Parent No. 2		Progeny phenotypes
	Genotype	Phenotype	Genotype	Phenotype	
	Genetic variation at structural locus; no genetic variation at modifier locus				
1.	*N/N ; m/m*	N	*N/N ; m/m*	N	N
2.	*N/V ; m/m*	N/V	*N/N ; m/m*	N	N, N/V
3.	*N/V ; m/m*	N/V	*N/V ; m/m*	N/V	N, N/V, V
	No genetic variation at structural locus; genetic variation at modifier locus				
4.	*N/N ; M/m*	(NM)	*N/N ; m/m*	N	(NM)*, N*
5.	*N/N ; M/m*	(NM)	*N/N ; M/m*	(NM)	(NM), N**
	Genetic variation at both structural and modifier loci				
6.	*N/V ; M/m*	(NM)/(VM)	*N/N ; m/m*	N	N*, N/V**, (NM)*, (NM)/(VM)***
7.	*N/V ; M/m*	(NM)/(VM)	*N/N ; M/m*	(NM)	N**, N/V**, (NM), (NM)/(VM)
8.	*N/V ; M/m*	(NM)/(VM)	*N/V ; M/m*	(NM)/(VM)	N**, N/V**, (NM), (NM)/(VM), (VM), V**
9.	*N/V ; m/m*	N/V	*N/N ; M/m*	(NM)	N*, N/V***, (NM)*, (NM)/(VM)**

N="normal" structural allele (co-dominant)
V="variant" structural allele (co-dominant)
m="non-modifier" allele at regulatory locus (recessive)
M="Modifier" allele at regulatory locus (dominant)
(NM) or (VM)=phenotypes resulting from modification of normal or variant structural gene product.
genotypes are given in italics; phenotypes given in non-italics

*—No heterozygous progeny observed for this purported "allele", suggesting dominant inheritance of phenotype.
**—appearance of progeny phenotype not present in either parent
***—apparent non-mendelian segregation in transmission of phenotype; no segregation in Parent No. 1.

alone) and matings involving an X-linked structural locus in association with autosomal variability for the modifier, will also produce unusual progeny (Table 1). The tables suggest that examination of family and population data should certainly be capable of providing positive evidence about the hypothesis that modifier variability acting through *GMPTM* is an important cause of electrophoretic variability, and a source of bias in the estimate of average heterozygosity in human populations.

3. MATERIALS AND METHODS

A complete review of the literature on genetic variation in humans is not feasible. Instead, we present a summary of the findings with respect to several informative loci, where the probability of detecting the effect of *GMPTM* is especially high. Three specific loci have been especially emphasized. These loci and the reasons for their selection are:

1) *Glucose-6-phosphate dehydrogenase* (*G6PD*)—this enzyme has been the subject of a large number of studies using a wide array of screening criteria in addition to standard electrophoresis. The X-linked nature of the gene also makes it easier to search for evidence of genetic segregation of a *G6PD* phenotype which would be incompatible with *G6PD* structural gene variation.

2) *Hemoglobin* (*HB*)—The products of the human α, β, δ, and γ hemoglobin loci have been the most intensively studied gene products in any higher organism. The involvement of multiple gene loci, with multiple variant alleles, and the availability of structural and functional studies on most of the reported genetic variants, together with the sheer volume of available data makes hemoglobin an ideal focus in the search for *GMPTM* effects.

3) *Alpha-1-antitrypsin* (*Pi*)—This locus has seen a startling increase in the reported number of alleles following the introduction of new screening procedures, especially isoelectric focusing. Data on both structural and biochemical properties exist for several of these alleles. *In vivo* posttranslational modification is involved in the production of the various isozymes which characterize the multibanded electrophoretic pattern obtained from the products of the locus. This system thus has a very high likelihood of yielding evidence of *GMPTM*, if it exists.

In addition to data for these three principal loci, other sources of information which have been examined include:

4) *Data for other loci which have variant alleles for which structural studies have been performed*—These include carbonic anhydrase-1 and phos-

phoglycerate kinase. Data from loci for which specific studies of post-translational modification have been conducted can also be examined.

5) *Data on the cotransmission of two traits associated with a specific presumptive variant allele.* These data will also include information on the frequency and family inheritance of null and activity alleles.

6) *Data from population surveys*—Population screening data can be used to address directly some of the "problems" cited by Johnson (1979) and Finnerty and Johnson (1979), and interpreted by them as being incompatible with simple structural gene variation. Among these problems are the "unexpectedly" high levels of genetic variation found when using new methods, and the apparent excess of rare alleles which is sometime discovered.

7) *Somatic cell hybridization studies*—The random loss of human chromosomes from human/rodent somatic cell hybrids provides a unique opportunity to search for the independent segregation of structural and modifier loci in cultured cell lines carrying various combinations of human chromosomes. Since almost all human polymorphic enzyme loci which are routinely typed in population surveys have been assigned to specific chromosomes by the somatic cell hybrid technique, there have been many opportunities to observe such independently segregating modifier loci, if they exist.

We are convinced that the data on human genetic variation presented here is representative of all available human data on biochemical electrophoretically detectable loci.

4. THE DATE

Glucose-6-Phosphate Dehydrogenase (G6PD)

Mammalian *G6PD* is a dimeric. X-linked enzyme, which functions physiologically to control the production of NADH (Ohno 1975; Yoshida 1973). Extensive data on over 200 distinct alleles have been collected using many different screening techniques for the genetic variation of the *G6PD* enzyme in man (Marti, Fischer and Killer 1977; Yoshida and Beutler 1983). Criteria used to classify variants often make use of a standard set of characteristics which include residual enzyme activity in erythrocytes, K_m's for glucose-6-phosphate and NADP, rate of utilization of the substrate analogue 2-deoxy-glucose-6-phosphate or the cofactor analogue deamino-NADP, heat stability, pH optimum, and electrophoretic migration under different electrophoretic conditions (Yoshida and Beutler 1983). Other methods which have been used to differentiate some variants include isoelectric focusing (Vergnes and Brun 1979), column chromatography (Usanga *et al.* 1977), and the use of multivariate analysis (Vergnes, Bonnet and Grozdea 1985).

Beutler (1978) tabulated data on over 120 genetic variants of human *G6PD*. Most of these variants are rare in all populations, but a few attain polymorphic frequencies in some populations. Of these 120 variants, 37 show mild to severe enzyme deficiency without being electrophoretically distinguishable from the normal *G6PD-B* allele product under the conditions used to study them. Thermal stability data are available for 104 *G6PD* variants; 78 variants differ in thermal stability from the *G6PD-B* allele. Alleles often share many properties, but most are uniquely identified when the entire set of physiochemical characters are considered. All of this variation has been ascribed to structural differences at the *G6PD* locus.

Table 1 shows possible progeny resulting from matings which involve allelic variants at an X-linked structural gene locus and alleles at a modifier locus on an autosomal chromosome. Matings #1 through #4 in Table 1 involve no modifier variability. They are expected to be observed in any potential test population, the presence of variation at a modifier locus obviously increases the number of possible genotypes expected in the progeny (see matings #5 to #12, Table 1). The presence of an autosomal locus which causes the post-translational modification of *G6PD*, and which possessed allelic variants that cause different modifications (or a locus at which some alleles cause modifications while others do not), would lead to segregation of *G6PD* variants in family studies in a manner that is incompatible with X-linkage. All matings in Table 1 which involve modifier heterozygosity produce some progeny which are incompatible with a single X-linked structural gene. Despite a large number of family studies on the genetic transmission of *G6PD*, no such cases have been reported (Giblett 1969; Beutler 1978). X-linked inheritance has invariably been observed for electrophoretic, stability, or activity variants of *G6PD*. Some allelic variants are different from the other alleles for both electrophoretic mobility and enzyme activity. For all such alleles, the two properties show concordant segregation when studied in families; no new properties were found in the progeny, such as should be seen for matings #5, #6, or #7 from Table I (Kahn *et al.* 1976b, c; Gherardi *et al.* 1976).

The interpretation that *GMPTM* does not contribute significantly to *G6PD* variation is further supported by independent segregation of multiple *G6PD* variants within the same family (Lenzerini *et al.* 1969; Vergnes *et al.* 1976; Usanga *et al.* 1977). These families represent potential examples of matings #8 or #9 from Table 1. Nevertheless, (with the possibility of a single exception, Gahr and Schrote 1975) no family has been reported to have discordant segregation, nor has there been any appearance in the progeny of these families of electrophoretic mobility, activity or thermal stability variation which was different from that observed in the parents. If trans-dominant post-translational modifier action were operating within this enzyme system, one would expect to observe some phenotypic evidence of independent assort-

ment of modifier and structural loci, exemplified in the progeny types of Table 1.

It is possible that differences in measured enzyme activity for *G6PD* might be caused by regulation of protein synthesis by a modifier locus. If this is so, the consequent interpretation of such activity differences as being caused by structural allele differences would be erroneous. This possibility has been tested for a series of eighteen variants by Kahn (1978) who concluded that reduced activity was related to either a reduced catalytic activity of the variant, or to increased lability of a variant which would lead to decreased activity in older erythrocytes. No evidence of decreased synthesis or of differential posttranslational modification was seen for this series of variants.

Studies by Kahn's group and others indicate, however, that a significant number of post-translational changes in the characteristics of *G6PD* do accompany cell aging (Dreyfus, Kahn and Schapira 1978; Kahn *et al.* 1974, 1976a; Rozenszajn *et al.* 1979; Fornaini *et al.* 1969). These alterations include the appearance of secondary isozymes which have isoelectric points different from the initial form. The appearance of these secondary isozymes accompanies changes in the kinetic properties of *G6PD* in older cells, as well as the changes in the ratio of enzymatic activity to immunological activity, suggesting that there is a partial inactivation of the enzyme. Electrophoretic properties of *G6PD* on starch gels are not significantly altered by aging (Fornaini *et al.* 1969). Treatment of *G6PD* with neuraminidase has no effect on the electrophoretic patterns, indicating that the enzyme is not subject to siaylations which could be misinterpreted as structural alterations (Kahn *et al.* 1974). None of these modifications of *G6PD* have been reported to be under genetic control. There has been a suggestion, however, that at least some of the modifications of *G6PD* in aging cells may be mediated by a peptide factor, and that this peptide factor may act during cell aging to cause changes in enzymes which are coded by other structural loci as well (Kahn *et al.* 1975). This is certainly the most suggestive evidence for the existence of any effect of *GMPTM* on *G6PD*.

Structural data identifying amino acid difference between different *G6PD* alleles has been slowly appearing. The available structural data on the variants *G6PD(A+)*, *G6PD(A−)*, *G6PD-Hektoen* and *G6PD-Manchester* show that each of these variants, which differ by many biochemical criteria from the normal *G6PD-B* allele, are truly the result of amino acid sequence changes caused by mutation at the *G6PD* structural locus (Yoshida 1977).

Finally, electrophoretic variants of dimeric enzymes usually produce a three banded pattern with relative staining intensity 1:2:1 in heterozygotes, due to the random aggregation of protein subunits. Activity variation associated with one of the subunits can often be detected as a distortion of this intensity pattern. In the case of *G6PD*, random X-chromosome inactivation in

heterozygous females leads to a double banded pattern, typical of that seen for monomeric enzymes. This occurs because heterodimers cannot be produced because of the cellular segregation of the inactivated copy of the two X-chromosomes into different cell lines. Electrophoretic variation produced by autosomal post-translational modifier genes could possibly lead to the reappearance of a triple (or even a four banded) heterozygous pattern in females. No such *G6PD* variation has been observed. It is possible, however, that because of the dominance relationships hypothesized for *GMPTM* (Finnerty and Johnson 1979), only one modifier phenotype would be present in a particular female, and this individual would be interpreted as having the rare completete inactivation of only one of the alleles in all cell lines.

Hemoglobin (Hb)

Hemoglobin is the product of the most extensively characterized set of structural genes in humans. In mammals, hemoglobin exists mainly as a tetramer, formed from the association of two dimers which in turn consist of two genetically distinct polypeptide chains, one alpha-like (α or ξ) chain, and one beta-like (β, δ, γ or ε) chain. With more than 300 human variants hemoglobins described, mutations have been associated with the four principle globin chain loci in the order $\beta > \alpha > \delta > \gamma$ (Lehmann and Kynoch 1976, Bunn, Forget and Ranney 1977). The hypothesis that "hidden" genetic variation which is not normally detected by standard electrophoretic techniques, but is revealed by such methods as gel sieving, isoelectric focusing, sequential electrophoretic screening, or other biochemical methods, will be enriched for variants which are the product of post-translational modification can be tested by using the extensive body of data available for the reported variants of human hemoglobin.

Lehmann and Kynoch (1976) tabulated some 150 variants of the alpha and beta chains of human hemoglobin. Of these variants, at least 30 are considered to be electrophoretically indistinguishable from normal hemoglobin A (*Hb-A*), the common phenotype, under "standard" electrophoretic conditions, usually cellulose acetate or starch gel electrophoresis in the range of pH 8.5–9.0. An additional 30 variants are indistinguishable from hemoglobin S (*Hb-S*), the next most frequently observed allelic type, and another 30 are listed as being the same as hemoglobin J (*Hb-J*). Considering only the results of electrophoresis under a single set of conditions, more than half of the variation at the alpha and beta hemoglobin loci is potentially "hidden" variation under normal screening conditions. We should note that the classification given by Lehmann and Kynoch (1976) represents a qualitative comparison, since many of these variants were never directly compared. However, recent studies by ourselves and others (Fuerst and Ferrell 1980, 1983; Ramshaw, Coyne and Lewontin

1979) indicate that a value of about 50% for the proportion of hemoglobin variation which is hidden under a single "standard" screening condition is reasonable.

Detailed structural analysis of the variants listed in Lehmann and Kynoch (1976) has shown that virtually all of this variation is due to single amino acid substitutions at one of the structural loci coding for the polypeptide chains of hemoglobin, the remainder of the variation is due to other mutational events in the structural genes which lead to error in chain termination, or are due to intra-cistronic crossing over, which is inferred to involve previously identified single amino acid mutations. None of the "hidden" variation revealed for hemoglobins can be attributed to post-translational modification. This conclusion holds not only for the electrephoretic variants of hemoglobin, but also for variants identified because of altered stability, such as the hemoglobins in the class *Hb-M*, or those variants which have altered oxygen affinity, including *Hb-Kansas*, *Hb-Koln* and *Hb-Yakima*, for example.

Much effort has been expended in attempts to develop multiple-condition screening criteria, which will uniquely identify all or most potential hemoglobin variation, and which would dispense with the necessity of identifying rare variants by means of amino acid sequencing (Schneider and Barwick 1978, 1979). These methods have included electrophoresis in various media, and at varying pH's, as well as the use of urea gel electrophoresis of the separated globin chains under different pH conditions· Polyacrylamide isoelectric focusing has also been added to the screening criteria used to define different hemoglobin alleles (Burdett and Whitehead 1977; Bassett *et al.* 1978). These methods have, in fact, proven capable of revealing most of the known "hidden" variants (Fuerst and Ferrell 1980; Ramshaw, Coyne and Lewontin 1979). A continued deficiency seems to exist among the reported variants for alleles which involve amino acid substitutions that do not cause changes in the net molecular charge of the molecule. This suggests that a further class of hidden structural variants may exist which resists identification by all the screening methods currently employed.

More important, however, for the purposes of the present argument is the fact that multiple screening methods have not produced any findings which suggest that *GMPTM* contributes to the previous or newly identified variability at the hemoglobin loci. All variants which have been detected during population screening efforts because of their unique migration patterns on multiple electrophoretic conditions have, thus far, always been shown to involve amino acid substitutions at one of the hemoglobin structural loci (Schneider *et al.* 1976).

A single survey produced data which may include some variation attributable to *GMPTM*. This survey investigated the use of heat denaturation to detect new hemoglobin variants (Bernstein, Bowman and Kaptue Noche 1980).

Several unique thermostability classes were shown to exist which had possible Mendelian segregation within small family units. Thermostability seemed to show concordant segregation with specific electrophoretic alleles, but thermostability classes greatly outnumbered electrophoretic classes. Whether the data represent structural differences (perhaps being a manifestation of the "missing" neutrally charged alleles) or is the result of *GMPTM* awaits further study.

A number of specific post-translational modification mechanisms do act on hemoglobin, and do result in shifts in the electrophoretic mobility of the molecule. Several of these have been suggested to be potentially important for *in vitro* studies, since they may cause electrophoretic artifacts by mechanisms such as the spontaneous interaction between hemoglobin and oxidized glutathione (Huisman and Dozy 1962) or by the enzymatic degradation of the alpha chain by carboxypeptidase B (Marti, Beale and Lehmann 1967; Kohne *et al.* 1977). Some of these mechanisms may also operate in aging cells, but if minimal care is taken in the collection of data, none of the mechanisms will result in the misclassification of electrophoretic variability, nor do any "variants" seem to have been inappropriately reported because of such mechanisms.

A more important group of modifications are those which occur *in vivo.* Reaction of the alpha-amino group of the normal beta chain with various hexoses leads to the production of electrophoretically and chromatographically distinct minor hemoglobins (Bunn *et al.* 1975). Because of the elevation of the quantity of these components in diabetics (Bunn, Gabbay and Gallop 1978), this phenomenon might be interpreted as a quantitative polymorphism in a population in which diabetes is especially frequent. However, these components do not segregate in a Mendelian fashion, and they have not apparently been confused with allelic variants in any population studies.

One additional group of modifications also occur *in vivo.* Several hemoglobin variants, each of which involve the substitution of asparagine for some other amino acid, sometime appear as two distinct electrophoretic types in a single individual, due to *in vivo* deamidation of asparagine to aspartic acid (Seid-Akhavan *et al.* 1976; Moo-Penn *et al.* 1976). The substitution of alanine for valine at the amino terminus of the beta chain has also been shown to result in modified mobility due to the acetylation of the alpha-amino group in the variant hemoglobin *Hb-Raleigh* (Moo-Penn *et al.* 1977). While only this later modification is known to be enzymatically mediated (Jornvall 1975), none of these post-translational modifications have been confused with genetic variability or genetic polymorphisms at a structural locus of hemoglobin.

For some time, a series of hemoglobinopathies called the thalassemias have been hypothesized to involve regulation of hemoglobin chain synthesis (Weatherall 1978). These diseases are characterized by a deficiency in the

amount of hemoglobin being produced by the erythrocyte. With the introduction of recombinant DNA techniques, it is now clear that these variants do not, in fact, involve the regulation of transcription or activity by genetic mechanisms. They are instead caused by either the deletion of a portion of the coding region of one of the globin genes, or involve a mutational change of a nucleotide which prevents the normal processing of introns in the globin precursor messenger-RNA (Weatherall and Clegg 1982).

Finally, we can again cite numerous family studies in which multiple variants are segregating in the same pedigree, and in which the phenotypes of the variant do not depend upon which other alleles are present in an individual or a parent (Steinberg *et al.* 1974; Tatsis, Sofronidou and Stergiopoulos 1976; Milner *et al.* 1976; Johnson, Powers and Schroeder 1976; Moo-Penn *et al.* 1977; to name but a few of many studies). Examination of Table 2, giving the expectations for an autosomal structural locus and an autosomal modifier locus suggests that if *GMPTM* were an important contributor to the variation reported for the hemoglobins, families representing types #4, #5, #6, or #9 should have been reported. To our knowledge, they have not.

Alpha-1-antitrypsin (*Pi*)

This serum protein is the major circulating inhibitor of serine proteases in human plasma (the designation *Pi* comes from protease inhibitor). An inherited deficiency state for alpha-1-antitrypsin has been shown to be related to the development of pulmonary emphysema and hepatic cirrhosis (Morse 1978). Several allelic variants at the *Pi* locus have been identified using acid starch electrophoresis combined with crossed agarose immunoelectrophoresis (Fagerhol and Laurell 1967). The protein appears as a series of six to eight bands under these conditions. Studies indicate that the various isozymes all have the same basic molecular size, with molecular weights of about 53,000 daltons (Yoshida *et al.* 1976; Yoshida and Wessels 1978). Recently, however, it has been found that minor differences between alleles may occur which reflect differences in the incorporation of neutrally charged sugars (Brown, 1982). The native molecule appears to function as a monomer.

The electrophoretic phenotype of *Pi* can be altered by treating the protein with neuraminidase, suggesting that the various *Pi* isozymes differ in the number of attached sialic acid residues (Headings and Bose 1974; Cox 1975a). Yoshida and Wessels (1978) presented evidence that the six major electrophoretic components of *Pi* differ by steps of a single sialic acid residue each. In the deficiency allele, *Pi-Z*, isozymes apparently lack one or two sialic acid residues compared to corresponding isozymes in the normal, *Pi-M*, allele (Cox 1975a). Developmental changes in the relative intensity of the different bands have been shown which suggest the maturation of an enzymatic system

which facilitates carbohydrate attachment to the protein core (Headings and Bose 1974).

Reports of new rare alleles at the locus increased dramatically following the introduction of isoelectric focusing methods to the study of *Pi* (Keuppers 1976, 1978; VandenBroek *et al.* 1976; Frants and Eriksson 1976; Arnaud *et al.* 1978a, b and others). In addition, much work has been expended upon quantitative studies of *Pi* levels in serum, which have domonstrated that most alleles show different quantitative patterns (Arnaud *et al.* 1978a; Fagerhol 1978; Kueppers 1978) and that quantitative alleles exist within the common electrophoretic phenotypes (Cox 1975b; Lie-Injo 1976; Keuppers, Utz and Simon 1977). All of this information suggests that the *Pi* system should be ideal for the search for evidence of *GMPTM* in man.

Despite this potential, *Pi* apparently provides no evidence to support the importance of genetically-mediated post-translational modification. The most promising candidate to be a major "allele" resulting from *GMPTM*, the deficiency *Pi-Z* allele, has been shown to differ from the predominant *Pi-M* allele by the substitution of a lysine for a glutamyl residue (Jeppson 1976; Yoshida *et al.* 1976). The change apparently blocks the normal sialylation of the protein and results in the difference in sialic acid content which was previously mentioned. The substitution causes both an alteration of the electrophoretic mobility of the protein, due to the amino acid substitution, and the alteration of the isozyme pattern, because of the lack of those isozyme bands which contain the greatest number of sialic acid residues.

The second most common allele of *Pi*, *Pi-S*, which is electrophoretically different from both *Pi-Z* and *Pi-M*, and which produces intermediate levels of the protein in serum, has been characterized by peptide mapping, and differs from the *Pi-M* allele by the substitution of valine for glutamic acid (Yoshida *et al.* 1977). This substitution has resulted in an electrophoretic shift of all bands, associated with the amino acid substitution. There is not, however, any accompanying change in the multi-band configuration of isozymes, because there is no change in sialylation (Owen and Carrell 1976).

Most impressive for the evidence of this paper, however, are studies concerning two alleles which were originally detected using isoelectric focusing. One allele, *Pi-B*^{Alhambra}, is a very rare anodic variant. Biochemical studies indicate that the allele possess normal trypsin inhibitory capacity (Yoshida, Chillar and Taylor 1979). Structural studies, using peptide analysis, show that the allele *Pi-B*^{Alhambra} differs from the standard *Pi-M*$_1$ allele by two amino acid substitutions, Lysine→Asparagine, and Glutamine→Asparagine, which occur in different peptides (Yoshida, Chillar and Taylor 1979). A second allele, *Pi-M*$_2$, is detectable only by using isoelectric focusing. It also has normal trypsin inhibitory capacity. This allele, however, is apparently common in most populations, reaching frequencies of almost 0.2 (Kueppers and Christo-

pherson 1978). Peptide analysis of the *Pi-M*$_2$ allele shows that it differs from the standard *Pi-M*$_1$ allele by a substitution of aspartic acid for glutamic acid in one residue (Yoshida, Taylor and VandenBroek 1979). This substitution, which involves no change in the net molecular charge of the molecule, was undetectable under the electrophoretic conditions used prior to the introduction of isoelectric focusing. The data thus show that all alleles which have been studied do differ by amino acid changes in the structural sequences of the protein.

The only suggestion of any abnormal segregation for *Pi* involves the apparent increase in the transmission of the *Pi-Z* allele for some heterozygotes (Chapuis-Cellier and Arnaud 1979; Iammarino, Wagener and Allen 1979). This has not, however, been seen in all studies (Cook 1975; Saugier *et al.* 1977; Constans, Chakraborty and Majunder 1983). The observation of abnormal segregation ratios may result from an undefined effect of the *Pi* deficiency in *Pi-Z*-bearing sperm, since the pattern of altered transmission seems to be seen only for heterozygous males (Chapuis-Cellier and Arnaud 1979; Iammarino *et al.* 1979). In this one instance, the possibility of some interaction between the *Pi* locus and other loci cannot be ruled out. Nevertheless, these interactions would not effect the heterozygosity levels estimated for the structural locus.

Examination of the data cited above for studies on the *Pi-Z* allele, however, demonstrate that there is no alteration of the electrophoretic phenotypes in families which would be consistent with the patterns predicted from the existence of modifier heterozygosity seen in Table 1. There has been no report of discordant transmission of trypsin inhibitory activity and electrophoretic type, nor any evidence that allele identification depends upon the presence of other alleles (Cook 1975; Kueppers 1978; VandenBroek *et al.* 1976; Cleve *et al.* 1979).

Other Specific Genetic Loci (Ca$_1$ *and PGK)*

Information which is less extensive, but similar in kind to the data summarized above for *G6PD*, *Hb* and *Pi* is available concerning the relationship between mobility, activity and structural differences for at least two other loci, carbonic anhydrase-1 (*Ca*$_1$) and Phosphoglycerate kinase (PGK).

Carbonic anhydrase is an enzyme responsible for maintaining correct serum pH in the serum. Several rare alleles exist at the *CA*$_1$ locus which have subtle but detectable electrophoretic differences (Tashian and Carter 1976). Several of these alleles also show altered activity or stability. Six of the variants have been studied structurally by amino acid sequencing, and compared with the *CA*$_1$*-1* allele. For each variant, differences from the normal allele could be attributed to a single nucleotide change which would cause an amino acid

substitution.

An analogous situation exists for the *PGK* locus, an enzyme in the glycolytic pathway. One variant allele was discovered which was indistinguishable from the *PGK* type allele under electrophoretic conditions which do not include citrate in the starch gel buffer (Yoshida *et al.* 1972). Structural studies indicate that this variant differs from the common allele, PGK^1, by a neutrally charged substitution (asparagine→threonine). Activity of the variant and the normal allele are identical. No data exist for either Ca_1 or *PGK* from families in which multiple rare variants are segregating. However, structural data suggest that *GMPTM* plays no role in the variation for the loci.

Mechanisms of post-translational modification for other human enzymes

The genetic systems reviewed above represent the most extensive body of data from humans for which inferences can be drawn concerning genetic control of post-translational modification. Geneticists have studied numerous other proteins in which post-translational modification could be a potential factor determining isozyme patterns seen using electrophoresis. We will review some of this data, evaluating it in light of the same questions which were used for the data from *G6PD*, *Hb*, and *Pi*; we can consider whether modifications of any sort occur and are detectable, whether activity or stability differences between presumed allelic variants have been reported, and finally whether information on primary structure or pedigree information implicate genetic factors other than structural gene variation in modifications which are observed.

What type of modifications occur?

Harris and Hopkinson (1977) summarized information for 130 enzyme loci. An examination of the descriptions of each enzyme system indicates that over half the loci show electrophoretic behavior under some conditions which could be the result of post-translational modification, and which might occur either *in vivo* or *in vitro*. For some loci, additional isozymes are clearly *in vitro* products which resulted from the conditions of storage of a sample. This effect is seen in systems such as isocitrate dehydrogenase, phosphogluconate dehydrogenase or the carbonic anhydrases. Other loci are known to be marked by an increased number of secondary isozymes which accompany tissue aging, an effect which is especially evident when erythrocytes are studied. Such changes are characteristic of nucleoside phosphorylase. Some enzymes show isozyme patterns which can be modified by neuraminidase treatment, suggesting the presence of sialic acid residues attached to the protein. Electrophoretic patterns for these loci are potentially modifiable *in vivo* by sialic acid cleavage, which could result in a change in molecular charge

without any alteration in the underlying amino acid sequence. Some of the loci falling into this category include the amylases and acid phosphatases. For none of these modifications has there been any evidence presented to indicate that genetic factors exist at significant frequencies which alter or mediate the post-translational changes which normally take place.

Activity variants and null alleles

Many loci are now being studied using methods which assess the enzymatic activity of different variants. The activity is a continuous trait, so that the definition of a variant with reduced or even "null" activity is somewhat arbitrary (Mohrenweiser 1983).

The "null" or "silent" allele class of alleles is particularly likely to include some variants which may be the product of post-translational modification. Such alleles are known for at least fourty of the loci reviewed by Harris and Hopkinson (1977), primarily because they are associated with a pathological phenotype. Because of their frequent pathological effect, the possibility for observing a homozygote excess (one of the predictions of *GMPTM*, according to Finnerty and Johnson 1979) is large. This is because the affected individual is usually identified through the health care system, because they suffer from ill health due to absence of the enzyme (because of homozygosity for the null allele), and are not being ascertained as part of a random population survey.

Electrophoretic mobility for silent alleles may be difficult to determine, because the residual activity of the enzyme is often so low that no bands appear following electrophoresis. On the other hand, considerable family information is often available because of the health problems posed by the allele. For over half the loci listed in Harris and Hopkinson (1977) at which silent alleles appear, family studies have been conducted which include the measurement of enzyme activity in relative of the affected putative homozygote. In all cases, individuals in the pedigree who should have been heterozygotes did exhibit activity levels which are intermediate between normal levels and the levels measured in the null homozygotes. These results suggest that the cause of the enzyme defect resides at the structural locus, and is not the result of a modification.

More recently, several surveys have been conducted specifically to identify deficiency variants for enzyme loci (Mohrenweiser 1981; Satoh *et al.* 1983; Mohrenweiser and Neel 1984). At all loci which were assayed, deficiency and activity variants were found. All fall into the class of rare alleles, with an average frequency of about 2-3/1000 gene products surveyed. In almost all cases, family studies confirmed that variant alleles which appeared in a child were also present in the parents.

One possible exception to this finding, however, is suggested by family studies of the pyruvate kinase (*PK*) deficiency in Japanese (Satoh *et al.* 1983; Miwa *et al.* 1980; Ishida *et al.* 1981). *PK* deficiency causes an anemia, which is associated with null activity in Caucasians in homozygotes, and no disease, but intermediate enzyme levels in heterozygotes (Valentine, Tanaka and Miwa 1961). In Japanese, at least seven different *PK* alleles with decreased activity have been identified (Miwa *et al.* 1980). These alleles retain some *PK* activity, in contrast to the null allele found in Caucasians, and consequently family studies often do not indicate significant reduction of *PK* activity in either parent of an affected individual. Affected individuals may, however, show quite high activity levels, suggesting heterozygosity for different variants may be the ultimate cause of the disease in many Japanese, not homozygosity for a null allele. Further work on this system to determine whether all low activity alleles are the result of the *PK* structural gene are necessary to resolve the issue satisfactorily. We should note that although these rare alleles have low activity, they do not seem to affect the thermostability of the enzyme (Satoh *et al.* 1985).

We should note one aspect of the hypothesis of *GMPTM* that applies to all of the null or activity alleles, and for the thermostability alleles to be discussed below, in human populations. This has to do with the problem of allele specific modification raised by Johnson and Finnerty (1979a). Since virtually all of these alleles have been detected in at least some individuals as heterozygotes, we must assume that there does exist structural differences between the alleles in an individual (since we assume that modifier allele interaction occurs in a dominant fashion). The only alternative is to assume that we are dealing with cis-acting modifiers. Only for the small number of enzymes in man in which two alleles are relatively common (such as acid phosphatase or phosphoglucomutase-1) could allele specific modification be considered a potential confounding factor.

In addition to null alleles, there are other allelic variants for which the alteration of enzyme activity is less severe. In addition to the studies of Mohrenweiser (1981) and Satoh *et al.* (1983) which looked at an array of enzymes, variants have been reported for pyridoxine kinase (Chern and Beutler 1976), inosine triphosphatase (Vanderheiden 1969), phosphoglucose isomerase (Satoh and Mohrenweiser 1979), and glutamic oxaloacetic transaminase (Wurzinger and Mohrenweiser 1982). In all of these studies, the inheritance of a variant without alteration of the activity pattern was confirmed in family studies. Finally, it should be noted that care must be taken in collecting data on enzyme activity, since there are known cases in which deficiency of a cofactor can cause an apparent loss of activity in a structural gene when none exists. This has been reported for glutatione reductase (Glatzle *et al.* 1974).

Stability variants

At several loci, variants have been reported which alter the thermostability or pH stability of the enzyme without having any major effect upon the electrophoretic mobility. We have already noted that this class of mutations occurs frequently for *G6PD*. Satoh and Mohrenweiser (1979) made a detailed study of the thermostability of variants for phosphoglucose isomerase which had been identified by activity criterion (see the last section), using a method of heat denaturation. Twenty-two heterozygous individuals for the normal allele (*GPI-1*) and an electromorph with altered activity (*GPI-4HIR*) were examined. The (*GPI-4HIR*) electrophoretic phenotype could be subdivided into three distinct thermal stability phenotypes. In all three classes, family studies showed that segregation of the electrophoretic phenotype was concordant with the thermal stability phenotype. Denaturation studies carried out in 5M urea confirmed the stability phenotypes, and the differences between the normal *GPI-1* class and the three subclasses of the *GPI-4HIR* electromorph. The most parsimonious interpretation of these results is that all observed variation is due to multiple alleles at the single structural locus for *GPI*.

These studies have been extended through population studies to other loci (Mohrenweiser and Neel 1981; Satoh *et al.* 1985). Seven different enzyme systems have been studied, and all systems have been found to contain thermostability variants. In every case where family studies could be performed, the results were consistent with a normal segregation of the thermostability variant, and not the segregation of unusual phenotypes predicted by Tables 1 and 2.

Positive evidence for GMPTM

Two loci may represent the only potential positive evidence in favor of a significant contribution of *GMPTM* to the genetic variability of a structural locus in human populations.

The reported polymorphism at the alpha-2-HS glycoprotein locus in human populations may represent such a case, since it seems to involve changes in non-protein moieties between alleles. This polymorphism is detected using two-dimensional electrophoresis (Anderson and Anderson 1979). The polymorphism consists of two alleles, the products of which clearly differ by carbohydrate content. It is not known what amino acid differences may exist between the alleles, so that a structural gene mutation has been neither confirmed nor ruled out.

A second case involves the interesting situation of the salivary amylases (AMY_1) reported by Karn *et al.* (1973). Two groups of isozymes attributable to amylase activity are detectable in saliva. Within each group, different

electrophoretic migration of isozymes can be produced by deamidation of the more cathodal forms. The two families of isozymes show molecular weight differences attributed to glycosidation. No other genetic loci have yet been implicated in causing these changes. A further modification, involving the deglycosidation of the previously glycosidated family of isozymes is clearly enzymatically mediated. However, the enzyme catalyzing this process is produced not by another locus in the human genome, but rather by the bacterial flora in the mouth. All isozymes can be ultimately traced back to the single product of one structural gene locus, and all presumptive alleles at this locus behave in a normal Mendelian manner within families, this system represents, however, a likely source of future evidence in favor of an effect of *GMPTM*.

5. POPULATION SURVEY DATA ON RARE ALLELES

Post-translational modification was suggested by Finnerty and Johnson (1979) and Johnson (1979) to explain a presumed excess of rare alleles, and the existence of these rare alleles as presumptive homozygotes in gel sieving studies inDrosophila.

The observation of "excess" rare alleles at protein loci in man is not, however, limited to techniques such as gel sieving, which are designed to reveal "hidden" electrophoretic variation during surveys. Chakraborty, Fuerst and Nei (1980) reported a detailed examination of the distribution of rare alleles in natural population which were identified using standard electrophoretic screening conditions. Human populations had the greatest excess of rare alleles among the 36 populations (26 species) of mammals examined. Humans were among the most deviant of all 138 populations examined in the study. If *GMPTM* is really an important source of such excess in rare alleles, we would expect evidence for this to be found by examining the rare alleles obtained in population surveys.

Homozygosity for rare alleles

From the arguments put forth by Johnson (1979) and Finnerty and Johnson (1979), rare alleles should appear most often as presumptive homozygotes, because of the dominance relationships at the modifier loci involved in *GMPTM*. A review of the available data on humans, summarized in Harris and Hopkinson (1977) has not revealed any instances of multiple homozygosity in the same individual for rare variants at different loci. Such cases would suggest that the products of several genes were being modified at the same time, perhaps due to alterations mediated by a gene coding for a glycosyl transferase, a glucosidase, or some other enzyme responsible for side chain degradation (O'Brien and Neufeld 1972). Alternatively, there could be a gene in man similar to that found by Finnerty, McCarron and Johnson (1979) which adds

some cofactor to different enzymes. If genetic variation in such modifying mechanisms were common, an excess of rare "homozygous" electrophoretic (or activity or stability) types would be expected (see Tables 1 and 2). Such an expectation is not supported for the human data. For virtually every locus listed in Harris and Hopkinson (1977) at which rare electrophoretic types have been reported, the rare allele was present in heterozygous combination with a common allele of the system. This makes the problem of allele specific modification as raised by Johnson and Finnerty (1979a) irrelevant, since heterozygosity for the rare electrophoretic type implies that there must be underlying structural differences between the alleles at the locus being modified (or else we are dealing with a cis-acting modifier).

At the population level, in a survey of genetic variation at 22 loci in the Japanese population, summarized by Neel *et al.* (1978), a total of 37 rare alleles were encountered in 174 individuals. All individuals showed apparent codominant expression of the alleles. In 85 cases, family studies confirmed the genetic character of the variants, while no case of distorted or aberrant segregation was encountered. Aberrant segregation would be expected in at least some cases if there were independent assortment of structural and modifier genes as detailed in Tables 1 and 2. At the transferrin locus, eight rare alleles were seen in 84 individuals. In matings between individuals heterozygous for different transferrin alleles, electrophoretic patterns expected for the inheritance of two codominant alleles at the transferrin structural locus were seen in all cases. At the phosphoglucomutase-1 locus, 8 rare alleles, in addition to the two common alleles were observed in 68 individuals. Again, all matings which involved different alleles showed inheritance patterns typical of codominant alleles at the structural gene locus, and did not show the patterns expected from Table 2 if modifier variation were present.

Other population studies involving twenty or more protein loci in thousands of individuals have revealed polymorphism at many loci, and the existence of rare alleles at virtually all loci without yielding and evidence that genetically mediated post-translational modification is responsible for the reported electrophoretic variation (Neel *et al.* 1978; Ferrell *et al.* 1978, 1979; studies summarized by Harris, Hopkinson and Robson 1973).

Johnson (1979) and Finnerty and Johnson (1980) have assumed that the overall level of variation seen for electrophoretic surveys is too large to be explained by most theories of population genetics. They assume that much "hidden" variability remains to be uncovered. As Neel (1984) has pointed out, in human populations this does not seem to be the case. We may already have uncovered between 65–75% of existing protein variation with one or two screening techniques. Instead the problem may be, as stated by Nei (1980) that variation is too low to be explained by the existing population genetic theory.

6. SOMATIC CELL HYBRIDIZATION STUDIES

Interspecific hybridization using somatic cells derived from man and rodents has led to the chromosomal assignment of a large number of human structural gene loci (Ruddle 1972; Shows and McAlpine 1979). While intrapopulation variation at loci mediating post-translational modification might be difficult to detect if it were at low frequency, somatic cell hybrids in which there is random segregation of human chromosomes in derivative clones should reveal any modifier loci which alter electrophoretic mobility at the non-syntenic primary structural gene products of interest. A review of the published data on gene assignments involving the electrophoretic analysis of somatic cell hybrids reveals that shifts in electrophoretic mobility caused by the hybridization of protein subunits between the rodent and human gene products in multimeric enzymes are easily detected (Shows 1972; Franke 1976). So, too, are shifts occuring because of the loss of a gene coding for one of the polypeptides in the cases of heteropolymeric enzymes (*LDH*, hexosaminidase, etc.). However, with the possible exception of the independent segregation of the structural locus for adenosine deaminase (*ADA*, found on chromosome 20) and a gene coding for the *ADA* complexing protein (found on chromosome 6), no independently segregating locus causing a post-translational modification seems to be present in these hybrids (Koch and Shows 1978). The existence of modification of *ADA* has not been confused with the typing of the structural gene polymorphism at the *ADA* locus.

7. OTHER CLASSES OF LOCI

The arguments presented in this paper are meant to apply only to the search for the effect of *GMPTM* on the estimated levels of electrophoretic variability for enzymes and soluble proteins in man. Virtually no information exists concerning levels of genetic variation of membrane bound proteins. Some data has been collected using two dimensional electrophoresis (Comings 1979; McConkey, Taylor and Phan 1979; Smith, Racine and Langley 1980), although the interpretation of the resolving power of 2-D electrophoresis remains somewhat questionable (Edwards and Hopkinson 1980; Warner, Neel and Meisler 1982; McLellan, Ames and Nikaido 1983). We can say nothing concerning possible effects of *GMPTM* on levels of variation estimated using these techniques.

Other genetic markers, especially immunologically detected systems such as the blood groups, may also differ from enzymes. Most blood groups seem to consists of membrane bound glycoproteins (Race and Sanger 1975). Antigenic differences are often conferred by differences in the attached oligosaccharide moiety (Watkins 1966). The relationship of these carbohydrate differences to underlying DNA structural variability has not been worked

out completely for most blood group systems. For instance, for the ABO blood group system, several different enzymes appear to be involved (see Race and Sanger 1975, pp. 59-63). The structural relationships among the various enzymes involved have not yet been elucidated, and the problem of explaining how enzymes with different specificities can appear to be allelic remains unresolved. Although it appears that the "allelic" variants of each blood group map to the same physical location, the possibility of *GMPTM* exists for such systems. However, since these loci have not normally been included in estimates of average genic heterozygosity in man (Harris and Hopkinson 1972; Nei and Roychoudhury 1974, 1982) and certainly have not been included in the estimates for other organisms, they do not alter our conclusions that *GMPTM* is not an important contributor to estimated genic variability in human populations.

8. DISCUSSION

Theoretically, the existence of sialic acid residues and other oligosaccharide moieties attached to many proteins provide ample opportunity for post-translational modification (Uy and Wold 1977). Many of these modifications would be enzymatically mediated. It is clear that in many cases, glycosyl transferases mediate the addition of a single specific sugar to an incomplete oligosaccharide side chain of the glycoprotein (O'Brien and Neufeld 1972). In other cases, it has been suggested that oligosaccharide chains are prefabricated and transferred intact to the protein acceptor, again presumably through the action of a transferase (Kennarz 1975). Other enzymes would, of course, then be involved in the fabrication of the oligosaccharide units. Irrespective of the method of synthesis, however, it is likely that the enzyme involved in the modification will specifically alter the products of more than a single genetic locus. In fact, the system presented by Finnnerty and Johnson (1979) in *Drosophila melanogaster* illustrates the pleiotrophic potential of modification, wherein the activities and stabilities of two genetic loci (*Xdh* and *AO*) were affected by the same modifier.

The existence of genetically determined, dominantly inherited post-translational modification of structural gene products has been observed in a number of vertebrate species. Law (1967) has shown that coordinate inheritance of electrophoretic variants of plasma alkaline phosphatase and leucine aminopeptidase in inbred lines of chickens was best explained by inheritance of an allele at a third locus controlling the attachment of sialic acid residues to both these enzymes. In the SM/J strain of the mouse, anodal migrating forms of three lysosomal enzymes are observed in the liver: acid phosphatase (Lalley and Shows 1977), alpha-mannosidase (Dizik and Elliott 1978) and acrylsulftase B (Daniel *et al.* 1979). Genes involved in the post-translational

modification of lysosomal enzymes have been implicated in each of these systems, and the variants at each of the loci have been mapped to approximately the same region of mouse chromosome 17 (Dizik and Elliott 1978; Daniel *et al.* 1979; Womack and Eichen 1977). The electrophoretic mobility of each of the enzymes in the SM/J strain can be converted to the pattern seen in other mouse strains by *in vitro* treatment of the enzymes with bacterial neuraminidase. Poitier *et al.* (1979) have demonstrated that in the SM/J mouse the liver is deficient in a specific component of neuraminidase and suggest that the deficiency leads to changes in post-translational processing of the lysosomal enzymes. The failure to cleave sialic acid residues from these proteins during or after transfer of the enzymes to the lysosomes accounts for the variant electrophoretic patterns.

Estimation of the levels of genetic variation in natural populations continues to be an important issue in population genetics. The recent uncovering of "hidden" variation at a variety of loci is not unexpected, since difference in electrophoretic mobility are largely determined by changes in the net molecular charge of a protein (Fuerst and Ferrell 1980) and only a fraction of the possible amino acid substitutions are expected to lead to changes in the net charge. Nei and Chakraborty (1976) have shown theoretically that the expected number of hidden alleles at a locus depends on population size, mutation rate and sample size. Their analyses assumed that only alleles with unit charge differences between them will be detected (the step-mutation model of Ohta and Kimura 1973), the actual degree of detectability for standard electrophoretic methods is somewhat greater than assumed by this model (Fuerst and Ferrell 1980), but it can be shown that many theoretical results are not qualitatively altered as long as the proportion of non-detectable substitutions is not very small (Li 1976). With the possible exception of the variation detected by gel sieving methods (Johnson 1975), which have not been extensively applied in population screening, the findings with respect to the level of "hidden" genetic variability in natural populations are not grossly different from what is theoretically expected (Nei and Chakraborty 1976; Selander and Whittam 1983).

While the functional significance of these levels of variability remain the subject of active research and debate, it seems clear that the genetic variation presently observed in human population results predominantly from single nucleotide substitutions in structural cistrons determining the amino acid sequence of the proteins being examined, and that *GMPTM* is neither required to explain anamolies of the patterns of variation, nor supported by any data. Ultimately, we must conclude that the studies which we have reviewed provide no evidence favoring an important contribution of *GMPTM* to the estimated levels of genic variability for enzymes or proteins in human populations.

9. SUMMARY AND CONCLUSIONS

1. Data on genetic variability for enzymes and proteins in humans were examined from several different perspectives concerning the hypothesis that dominantly inherited genetically mediated post-translational modifiers (*GMPTM*) contribute significantly to the measured levels of genic heterozygosity in man.

2. Three proteins, Glucose-6-phosphate dehydrogenase, Hemoglobin and Alpha-1-antitrypsin were extensively examined because of the large amount of information available from studies on these proteins. Other loci were briefly reviewed. No evidence was uncovered suggesting that *GMPTM* has contributed to previous estimates of genetic variability at these proteins, or that techniques of uncovering "hidden" variability have been misleading by attributing all variant alleles to structural modifications coded directly by the locus under consideration.

3. The familial inheritance patterns expected when both structural and modifier variability exist in the population are given (Tables 1 and 2). It is concluded that is genetic variation is moderately frequent for both types of loci, aberrant Mendelian segregation should be easy to observe. Such unusual family material is notable only by its absence.

4. When multiple characteristics of rare alleles are examined, such as biochemical activity, electrophoretic mobility under various conditions, molecular stability, etc., family data invariably demonstrates the co-segregation of properties. In contrast to the assertion of proponents of *GMPTM*, there is no apparent excess of homozygotes for rare alleles in human, even when alleles are identified using non-standard screening methods.

5. When the structural basis of variability has been determined (especially for data on hemoglobins) the variant phenotype, either electrophoretic or functional, is always explained by amino acid differences between alleles.

6. Data from population surveys in man and from human/rodent somatic cell hybrids also fail to suggest an important role for *GMPTM*.

7. Finally, post-translational mechanisms do affect each of the loci reviewed. Although electrophoretic phenotypes are often altered during cell aging, sample storage, etc., these changes are not genetically mediated, and there is no evidence that measures of genic variability have been biased because of *GMPTM*.

REFERENCES

ANDERSON, N. L. and ANDERSON, N. G. (1979) Microheterogeneity of serum transferrin, haptoglobin, and alpha-2-HS glycoprotein examined by high resolution two-dimensional electrophoresis. *Biochem. Biophys. Res. Comm.* **88**, 258–265.

ARNAUD, P., CHAPUSI-CELLIER, C., VITTOZ, P. and FUDENBERG, H. H. (1978a) Genetic polymorphism of serum alpha-1-protease inhibitor (alpha-1-antitrypsin): Pi I, a deficient allele of the Pi systems. *J. Lab. Clin. Med.* **92**, 177–184.

ARNAUD, P., GALBRAITH, R. M., GALBRAITH, G. M. P., ALLEN, R. C. and FUDENBERG, R. R. (1978b) A new allele of human alpha-1-antitrypsin: $Pi^{Nhampton}$. *Amer. J. Hum. Genet.* **30**, 653–659.

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

BASSET, P., BENZARD, Y., GAREL, M. C. and ROSA, J. (1978) Isoelectric focusing of human hemoglobins: its application to screening, to the characterization of 70 variants, and to the study of modified fractions of normal hemoglobins. *Blood* **51**, 971–982.

BERNSTEIN, S. C., BOWMAN, J. E. and KAPTUE NOCHE, L. (1980) Genetic variation in Cameroon: thermostability variants of hemoglobin and of glucose-6-phosphate dehydrogenase. *Biochem. Genet.* **18**, 21–37.

BEUTLER, E. (1978) Hemolytic Anemia in Disorders of Red Cell Metabolism. Plenum, New York.

BROWN, W. T. (1982) Alpha-1-antitrypsin: apparent molecular weight heterogeneity shown by two-dimensional electrophoresis. *Amer. J. Hum. Genet.* **34**, 195–208.

BUNN, H. F., FORGET, B. G. and RANNEY, H. M. (1977) *Human Hemoglobins.* Saunders Co., New York.

BUNN, H. F., GABBAY, K. H. and GALLOP, P. M. (1978) The glycosylation of hemoglobin: Relevance to diabetes mellitus. *Science* **170**, 21–27.

BUNN, H. F., HANEY, D. H., GABBAY, K. H. and GALLOP, P. M. (1975) Further identification of the nature and linkage of the carbohydrate in hemoglobin A_{1c}. *Biochem. Biophys. Res. Comm.* **678**, 103–109.

BURDETT, P. E. and WHITEHEAD, P. H. (1977) The separation of the phenotypes of phosphoglucomutase, erythrocyte acid phosphatase and some haemoglobin variants using isoelectric focusing. *Anal. Biochem.* **77**, 419–428.

CHAKRABORTY, R., FUERST, P. A. and NEI, M. (1980) Statistical studies on protein polymorphism in natural populations· III. Distribution of allele frequencies and the number of alleles per locus. *Genetics* **94**, 1039–1063.

CHAPUIS-CELLIER, C. and ARNAUD, P. (1979) Preferential transmission of the Z deficient allele of alpha-1-antitrypsin. *Science* **205**, 407–408.

CHERN, C. J. and BEUTLER, E. (1976) Biochemical and electrophoretic studies of erythrocyte pyridoxine kinase in white and black Americans. *Amer. J. Hum. Genet.* **28**, 9–17.

CLEVE, H., PATUTSCHNICK, W., STRECKER, K. and NEVO, S. (1979) Inheritance of PiM subtypes: A study of 151 families with a total of 242 children and 142 mother-child pairs. *Hum. Hered.* **29**, 351–354.

COMINGS, D. E. (1979) Pc 1 Duarte, a common polymorpism of a human brain protein, and its relationship to depressive disease and multiple sclerosis. *Nature* **277**, 28–32.

CONSTANS, J., CHAKRABORTY, R. and MAJUNDER, P. P. (1982) Transmission of Z allele from heterozygote males for alpha-1-antitrypsin deficiency: additional family data. *Amer. J. Hum. Genet.* **34**, 674–675.

COOK, P. J. L. (1975) The genetics of alpha-1-antitrypsin: a family study in England and Scotland. *Ann. Hum. Genet.* **30**, 275–287.

COX, D. W. (1975a) The effect of neuraminidase on genetic variants of alpha-1-antitrypsin. *Amer. J. Hum. Genet.* **28**, 165–177.

COX, D. W. (1975b) Deficiency allele of alpha-1-antitrypsin: PiM^{malton}. *Amer. J. Hum. Genet.* **27**, 29A.

COYNE, J. A., EANES, W. F. and LEWONTIN, R. C. (1979) The genetics of electrophoretic variation. *Genetics* **92**, 358–361.

DANIEL, W. L., WOMACK, J. E. and HENTHORN, P. S. (1979) Murine liver arylsulfatase B processing gene located on chromosome 17. *Genetics.*

DIZIK, M. and ELLIOTT, R. W. (1978) A second gene affecting the sialylation of lysosomal alpha-mannosidase in mouse liver. *Biochem. Genet.* **16**, 247–260.

DREYFUS, J. C., KAHN, A. and SCHAPIRA, F. (1978) Posttranslational modification of enzymes. *Curr. Topics Cellular Rispiration* **14**, 243–297.

DYKHUISEN, D. E., MUDD, C., HONEYCUTT, A. and HARTL, D. L. (1985) Polymorphic posttranslational modification of alkaline phosphatase in *Escherichia coli*. *Evolution* **39**, 1–7.

EDWARDS, Y. H. and HOPKINSON, D. A. (1980) Are abundant proteins less variable? *Nature* **284**, 511–512.

FAGERHOL, M. K. and LAURELL, G. B. (1967) The polymorphism of "prealbumins" and alpha-1-antitrypsin in human sera. *Clin. Chim. Acta* **16**, 199-203.

FAGERHOL, M. K. (1968) Quantitative studies on the inherited variants of serum alpha-1-antitrypsin. *Scand. J. Clin. Invest.* **23**, 97-103.

FERRELL, R. E., BERTIN, T., BARTON, S. A., ROTHHAMMER, F. and SCHULL, W. J. (1979) The Multinational Andean Genetic and Health Program. IX. Gene frequencies and rare variants of 20 serum proteins and erythrocyte enzymes in the Aymara of Chile. *Amer. J. Hum. Genet.* (in press).

FERRELL, R. E., BERTIN, T., YOUNG, R., BARTON, S. A., MURILLO, F. and SCHULL, W. J. (1978) The Aymara of Western Bolivia. IV. Gene frequencies for eight blood groups and 19 protein and erythrocyte enzyme systems. *Am. J. Hum. Genet.* **30**, 539-549.

FINNERTY, V. and JOHNSON, G. (1979) Post-translational modification as a potential explanation of high levels of enzyme polymorphism: Xanthine dehydrogenase and aldehyde oxidase in *Drosophila melanogaster*. *Genetics* **91**, 695-722.

FINNERTY, V., MCCARRON, M. and JOHNSON, G. B. (1979) Gene expression in *Drosophila*: Post-translational modification of aldehyde oxidase and xanthine dehydrogenase. *Molec. Genet.* **172**, 37-43.

FORNAINI, G., LEONCICI, G., SEGNI, P., CALABRIA, G. A. and DACHA, M. (1969) Relationship between age and properties of human and rabbit erythrocyte glucose-6-phosphate dehydrogenase. *Eur. J. Biochem.* **7**, 214-222.

FRANKE, U. (1976) The human gene for Beta-glucuronidase is on chromosome 7. *Amer. J. Hum. Genet.* **28**, 357.

FRANTS, R. R. and ERIKSSON, A. W. (1978) Reliable classification of six PiM subtypes by separator isoelectric focusing. *Hum. Hered.* **28**, 201-209.

FUERST, P. A. and FERRELL, R. E. (1980) The stepwise mutation model: an experimental evaluation utilizing hemoglobin variants. *Genetics* **94**, 185-201.

FUERST, P. A. and FERRELL, R. E. (1983) The analysis of hidden electrophoretic variation: Interspecific electrophoretic differentiation and amino acid divergence. *J. Mol. Evol.* **19**, 449-454.

GAHR, M. and SCHROTER, W. (1975) Inhibition of normal red cell glucose-6-phosphate dehydrogenase (G-6-PD) by a variant enzyme, G-6-PD Hamburg. *FEBS Letters* **59**, 237-40.

GHERARDI, M., BIERME, R., CORBERAND, J., PRIS, J. and VERGNES, H. (1976) Distribution of G6PD types in the population of southwest France: Common variants and new variants. *Hum. Hered.* **26**, 279-289.

GIBLETT, E. R. (1976) *Genetic Markers in Human Blood*. Blackwell, Oxford.

GLATZLE, D., VUILLEUMIER, J. P., WEBER, F. and DEKKER, K. (1974) *Experientia* **30**, 665-667.

HARRIS, H. (1969) Genes and isozymes. *Proc. Roy. Soc. B* **174**, 1-10.

HARRIS, H. and HOPKINSON, D. A. (1972) Average heterozygosity per locus in man: an estimate based on the incidence of enzyme polymorphism. *Ann. Hum. Genet.* **36**, 9-20.

HARRIS, H. and HOPKINSON, D. A. (1977) *Handbook of Enzyme Electrophoresis in Human Genetics*. North-Holland-American Elsevier, New York.

HARRIS, H., HOPKINSON, D. A. and ROBSON, E. B. (1973) The incidence of rare alleles determining electrophoretic variants: Data on 43 enzymes in man. *Ann. Hum. Genet.* **37**, 237-253.

HEADINGS, V. E. and BOSE, S. (1974) Multiple molecular forms of alpha-1-antitrypsin. *Biochem. Genet.* **11**, 429-440.

HOPKINSON, D. A. and HARRIS, H. (1971) Recent work in isozymes in man. *Ann. Rev. Genet.* **5**, 5-32.

HUISMAN, T. H. J. and DOZY, A. M. (1962) Studies on the heterogeneity of hemoglobin. V. Binding of hemoglobin with oxidized glutathione. *J. Lab. Clin. Med.* **60**, 302-319.

IAMMARINO, R. M., WAGENER, D. K. and ALLEN, R. C. (1979) Segregation distortion of the alpha-1-antitrypsin PiZ allele. *Amer. J. Hum. Genet.* **31**, 508-517.

ISHIDA, Y., MIWA, S., FUJII, H., FUJINAMI, N., TAKEGAWA, S. and YAMATO, K. (1981) Trirteen cases of pyruvate kinase deficiency found in Japan. *Amer. J. Hematol.* **10**, 239-250.

JEPPSON, J. O. (1976) Amino acid substitution Glu→Lys in alpha-1-antitrypsin Pi Z. *FEBS Letters* **65**, 1195-1197.

JOHNSON, C., POWERS, D. and SCHROEDER, W. A. (1976) A case of both hemoglobins C and N-Baltimor. *Acta Haemat.* **56**, 83-183.

JOHNSON, G. B. (1975) Use of internal standards in electrophoretic surveys of enzyme polymorphism. *Biochem. Genet.* **13**, 833-847.

JOHNSON, G. B. (1979) Genetically controlled variation in the shapes of enzymes. *Progr. Nucl. Acid Res. Molec. Biol.* **22**, 293-426.

JOHNSON, G. B. and FINNERTY, V. (1979a) The genetics of electrophoretic variation. A reply. *Genetics* **92**, 357-360.

JOHNSON, G. B. and FINNERTY, V. (1979b) Structural vs. post-translational components of genic variation. A reply. *Genetics* **92**, 683-684.

JOHNSON, G., FINNERTY, V. and HARTL, D. (1981) Post-translational modification of xanthine dehydrogenase in a natural population of *Drosophila melanogaster*. *Genetics* **98**, 817-831.

JORNVALL, H. (1975) Acetylation of protein N-terminal amino groups: Structural observations on alpha-amino acetylated proteins. *J. Theor. Biol.* **55**, 1-12.

KAHN, A. (1978) G6PD variants. *Hum. Genet. Suppl.* **1**, 37-44.

KAHN, A., BOIVIN, P., VIBERT, M., COTTREAU, D. and DREYFUS, J. C. (1974) Post-translational modifications of human glucose-6-phosphate dehydrogenase. *Biochimie* **56**, 1395-1407.

KAHN, A., BERTRAND, O., COTTREAU, D., BOIVIN, P. and DREYFUSS, J. C. (1976a) Studies on the nature of glucose-6-phosphate dehydrogenase purified from human leukocytes. *Biochim. Biophys. Acta* **445**, 537-548.

KAHN, A., BOIVIN, P., RUBINSON, H., COTTREAU, D., MARIE, J. and DREYFUS, J. C. (1976) Modification of purified glucose-6-phosphate sehydrogenase and other enzymes by a factor of low molecular weight abundant in some leukemic cells. *Proc. Nat. Acad. Sci. USA* **73**, 77-81.

KAHN, A., ESTERS, A. and HABEDANK, M. (1976b) Gd(−) Aachen, a new variant of deficient glucose-6-phosphate dehydrogenase. *Hum. Genet.* **32**, 171-180.

KAHN, A., MARIE, J., DESBOIS, J.C. and BOIVIN, P. (1976c) Favism in a Portuguese family due to a deficient glucose-6-phosphate dehydrogenase variant of "Canton" or "Canton-like" type. *Acta Haematol.* **56**, 58-64.

KARN, R. C., SHULKIN, J. O., MERRITT, A. D. and NEWELL, R. C. (1973) Evidence for posttranscriptional modification of human salivary amylase (Amy_1) isozymes. *Biochem Genet.* **10**, 341-350.

KOCH, G. and SHOWS, T. B. (1978) A gene on human chromosome 6 functions in assembly of tissue-specific adenosine deaminase isoenzymes. *Proc. Nat. Acad. Sci. (USA)* **75**, 3876-3880.

KOHNE, E., KRAUSE, M., LEOPOLD, D. and KLEINHAUER, E. (1977) Hemoglobin F Koelliker (α_2 minus 141 (HC_3) Arg β_2): A modification of fetal hemoglebin. *Hemoglobin* **1**, 257-266.

KUEPPERS, F. (1976) Alpha-1-antitrypsin M_1: A new common genetically determined variant. *Amer. J. Hum. Genet.* **28**, 270-277.

KUEPPERS, F. and CHRISTOPHERSON, M. J. (1978) Alpha-1-antitrypsin: Further genetic heterogeneity revealed by isoelectric focusing. *Amer. J. Hum. Genet.* **30**, 359-365.

KUEPPERS, F., UTZ, G. and SIMON, B. (1977) Alpha-1-antitrypsin deficiency with M-like phenotype. *J. Med. Genet.* **14**, 183-186.

KUEPPERS, F. (1978) Inherited differences in alpha-1-antitrypsin. pp. 23-74. In: S. Litwin (ed.), *Lung Biology in Health and Disease*. Vol. 11: Genetic determinants of pulmonary diseases. Marcel Dekker, New York.

LALLEY, P. A. and SHOWS, T. B. (1977) Lysosomal acid phosphatase deficiency: liver specific variant in the mouse. *Genetics* **87**, 305-317.

LAW, G. R. J. (1967) Alkaline phosphatase and leucine aminopeptidase association in plasma of the chicken. *Science* **156**, 1106-1107.

LEHMANN, H. and KYNOCH, P. A. (1976) Human Hemoglobin Variants and Their Characteristics. North-Holland, Amsterdam.

LENNARZ, W. J. (1975) Lipid linked sugars in glycoprotein synthesis. *Science* **188**, 986-991.

LENZERINI, L., MEERA KHAN, P., FILIPPI, G., RATTAZZI, M. C., RAY, A. K. and SINISCALCO, M. (1969) Characterization of glucose-67-phosphate dehydrogenase variants. 1. Occurrence of G6PD Seattle-like variant in Sardinia and its interaction with the G6PD Mediterranean variant. *Amer. J. Hum. Genet.* **21**, 142-153.

LI, W. H. (1976) A mixed model of mutation for electrophoretic identity of proteins within and between populations. *Genetics* **83**, 423-432.

LIE-INJO, L. E. (.976) Alpha-1-antitrypsin with unusual behavior. *Clin. Chim. Acta* **82**, 83-87.

MARTI, H. R., BEALE, D. and LEHMANN, H. (1967) Haemoglobin Koelliker: A new acquired haemoglobin appearing after severe haemolysis: α_2 minus 141 Arg β_2. *Acta Haemat.* **37**, 174-180.

MARTI, H. R., FISCHER, S. and KILLER, D. (1977) Characterization of abnormal glucose-6-phosphate dehydrogenase variants. *Acta Biol. Med. Germ.* **36**, 703-708.

MCLELLAN, T., AMES, G. F-L. and NIKAIDO, K. (1983) Genetic variation in proteins: comparisons of one-dimensional and two-dimensional electrophoresis. *Genetics* **104**, 381-390.

MCCONKEY, E.H., TAYLOR, B. J. and PHAN, D. (1979) Human heterozygosity: a new estimate. *Proc. Nat. Acad. Sci. USA* **76**, 6500-6504.

MILNER, P. F., CORLEY, C. C., POMEROY, W. L., WILSON, J. B., GRAVELY, M. and HUISMAN, T. H. J. (1976) Thalassemia intermedia caused by both betathalassemia and hemoglobin Saki ($\beta_{14}(A_{11})$Leu). *Proc. Am. J. Haemat.* **1**, 283-292.

MIWA, S., FUJII, H., TAKEGAWA, S., ISHIDA, Y. FUJINAMI, N. and YAMATO, K. (1980) Seven pyruvate kinase variants characterized by the ISCH recommended methods. *Brit. J. Haematol.* **45**, 575-583.

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

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

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

MOHRENWEISER, H. W. and NEEL, J. V. (1984) A "disproportion" between the frequency of rare electromorphs and enzyme deficiency variants in Amerindians. *Amer. Jour. Hum. Genet.* **36**, 655-662.

MOO-PENN, W. F., JUE, D. L., BECHTEL, K. C., JOHNSON, M. H. and SCHMIDT, R. M. (1976) Hemoglobin Providence. A human hemoglobin variant occurring in two forms *in vivo*. *J. Biol. Chem.* **251**, 7557-7562.

MOO-PENN, W. F., BECHTEL, K. C., SCHMIDT, R. M., JOHNSON, M. R., JUE, D. L., SCHMIDT, D. E., DUNLAP, W. M., OPELLA, S. J., BONAVENTURA, J. and BONAVENTURA, C. (1977) Hemoglobin Raleigh (β_1 valine→Acetylalamine). Structural and functional chracterization. *Biochem.* **16**, 4872-4789.

MORSE, J. O. (1978) Alpha-1-antitrypsin deficiency. *New Engl. J. Med.* **299**, 1045-1048, 1099-1105.

NEEL, J. V., UEDA, N., SATOH, G., FERRELL, R. E., TANIS, R. J. and HAMILTON, R. 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**, 428-441.

NEEL, J. V. (1984) A revised estimate of the amount of genetic variation in human proteins: implications for the distribution of DNA polymorphisms. *Amer. J. Hum. Genet.* **36**, 1135-1148.

NEI, M. (1980) Stochastic theory of population genetics and evolution. pp. 17-47. In: Barigozzi, C. (ed.) *Vito Volterra Symposium on Mathematical Models in Biology*. Springer-Verlag, Berlin.

NEI, M. and CHAKRABORTY, (1976) Electrophoretically silent alleles in a finite population. *J. Mol. Evol.* **8**, 381-835.

NEI, M. and ROYCHODHURY, A. K. (1974) Genetic variation within and between the three major races of man, Caucasoids, Negroids, and Mongoloids. *Amer. J. Hum. Genet.* **26**, 421-443.

NEI, M. and ROYCHOUDHURY, A. K. (1982) Genetic relationship and evolution of human races. *Evolutionary Biology* **14**, 1-59.

O'BRIEN, P. J. and NEUFELD, E. F. (1972) Biosynthesis of Glycoproteins. Glycosyltransferases. In: Gottschalk, A. (ed.) Glycoproteins p. 1170-1186. Elsevier, New York.

OHNO, S. (1975) Ancient linkage groups and frozen accidents. *Nature* **244**, 259-262.

OHTA, T. and KIMURA, M. (1973) A model of mutation appropriate to estimate the number of electrophoretically detectable alleles in a genetic population. *Genet. Res.* **22**, 201-204.

OWEN, M. C. and CARRELL, R. W. (1976) Alpha-1-antitrypsin: Molecular abnormality of S variant. *Brit. Med. J.* **1**, 130-131.

POTIER, M., YAN, D. L. S. and WOMACK, J. E. (1979) Neuraminidase deficiency in the mouse. FEBS letters

RACE, R. R. and SANGER, R. (1975) Blood groups in man. Sixth edition. Blackwell Scientific Publ. Oxford.

RAMSHAW, J. A. M., COYNE, J. A. and LEWONTIN, R. C. (1979) The sensitivity of gel electrophoresis as a detector of genetic variation. *Genetics* **93**, 1010–1037.

ROZENSZAJN, L. A., KOLMAN, S., SHOHAM, D. and GAZITH, J. (1970) Glucose-6-phosphate dehydrogenase isoenzymes in blood cells. *Nature* **226**, 862–863.

RUDDLE, F. H. (1972) Linkage analysis using somatic cell hybrids. In: Harris, H. and K. Hirschhorn (eds.) Advances in Human Genetics. Vol. 3, p. 173–205. Plenum, New York.

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

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

SATOH, C., NEEL, J. V., MIURA, A., UENO, C., ARAKAWA, H., OMINE, H., GORIKI, K. and FUJITA, M. (1985) Inherited thermostability variants of seven enzymes in a Japanese population. *Ann. Hum. Genet.* **49**, 11–22.

SAUGIER, B., CHAPUIS-CELLIER, C., EUMONOT, A., VITTOZ, P. and GALY, P. (1977) Deficit en alpha-1-antitrypsine. Etude du phenotype de 60 membres d'une meme famille. *Sem. Hop. Paris* **53**, 413–416.

SCHNEIDER, R. G. and BARWICK, R. (1978) Measuring relative electrophoretic mobilities of mutant hemoglobins and globin chains. *Hemoglobin* **2**, 417–435.

SCHNEIDER, R. G. and BARWICK, R. C. (1979) Section on hemoglobins. In: R. M. Schmidt (ed.) *CRC Handbook in Clinical Laboratory Sciences*, Vol. 1, Hematology. Chemical Rubber Co., Cleveland, Ohio.

SCHNEIDER, R. G., HIGHTOWER, B., HOSTY, T. S., RYDER, H., TAMLIN, G., ATKINS, R., BRIMHALL, B. and JONES, R. T. (1976) Abnormal hemoglobins in a quarter million people. *Blood* **48**, 629–637.

SEID-AKHAVAN, M., WINTER, W. P., ABRAMSON, R. K. and RUCKNAGEL, D. L. (1976) Hemoglobin Wayne: A frameshift mutation detected in human hemoglobin alpha chains. *Proc. Nat. Acad. Sci. (USA)* **73**, 882–886.

SELANDER, R. K. and WHITTAM, T. S. (1983) Protein polymorphism and the genetic structure of populations. In: Nei, M. and R. K. Koehn (eds). Evolution of Genes and Proteins. p. 89–114. Sinamer, Sunderland, Mass.

SHOWS, T. B. (1972) Linkage of loci for human pyruvate kinase and mannose phosphatase isomerase in somatic cell hybrids. *Amer. J. Hum. Genet.* **24**, 13a.

SHOWS, T. B. and MCALPINE, P. J. (1979) The 1979 catalog of human genes and chromosome assignments. *Cytogenet. Cell Genet.* **25**, 117–127.

SINGH, R. S. (1979) Structural vs. post-translational components of gene variation. *Genetics* **92**, 679–682.

SMITH, S. C., RACINE, R. R. and LANGLEY, C. H. (1980) Lack of gene variation in the abundant proteins of human kidney. *Genetics* **98**, 179–197.

STEINBERG, M. H., ADAMS, J. G., THIGPEN, J. T., MORRISON, F. S. and DREILING, B. J. (1974) Hemoglobin Hope ($\alpha_2\beta_2$ 136-Gly→Asp)-S disease: Clincial and biochemical studies. *J. Lab. Clin. Med.* **84**, 632–642.

TASHIAN, R. E. and CARTER, N. D. (1976) Biochemical genetics of carbonic anhydrase. In: Harris, H. and K. Hirschhorn (eds.) Advances in Human Genetics, Vol. 7, p. 1–56. Plenum, New York.

TATSIS, B., SOFRONIADOU, K. and STERGIOPOULOS, C. I. (1976) Hemoglobin Pyrgos $\alpha_2\beta_2$ 83(EF7) Gly→Asp: A new hemoglobin variant in double heterozygosity with hemoglobin S. *Blood* **47**, 827–832.

TURNER, B. C., FISHER, R. A. and HARRIS, H. (1975) Posttranslational alterations of human erythrocyte enzymes. In: C. L. Markert (ed.), *Isozymes*, Vol. 1, pp. 781–795. Academic

Press, New York.

USANGA, E. A., BIENZLE, U., CANCEDDA, R., FASUAN, F. A., AJAYI, O. and LUZZATTO, L. (1977) Genetic variants of human erythrocyte glucose-6-phosphate dehydrogenase: New variants in West Africa characterized by column chromatography. *Ann. Hum. Genet.* **40**, 279-286.

UY, R. and WOLD, F. (1977) Post-translational covalent modification of proteins. *Science* **198**, 890-896.

VALENTINE, W. N., TANAKA, K. R. and MIWA, S. (1961) A specific erythrocyte glycolytic enzyme defect (pyruvate kinase) in three subjects with congenital non-spherocytic hemolytic anemia. *Trans. Assoc. Amer. Physicians* **74**, 100-110.

VANDENBROEK, W. G. M., HOFFMAN, J. J. M. L. and DIJKMAN, (1976) A new, high frequency variant of alpha-1-autitrypsin. *Hum. Genet.* **34**, 17-22.

VANDERHEIDEN, B. S. (1969) Genetic studies of human erythrocyte inosine triphosphate. *Biochem. Genet.* **3**, 289-295.

VERGNES, H. A., BONNET, L. G. and GROZDEA, J. D. (1985) Genetic variants of erythrocyte glucose-6-phosphate dehydrogenase: new characterization data obtained by multivariate analysis. *Ann. Hum. Genet.* **49**, 1-9.

VERGNES,H. A. and BRUN, H. (1979) Characterization of some erythrocyte G6PD variants by isoelectric focusing. *Hum. Genet.* **53**, 43-45.

VERGNES, H., GHERARDI, M. and YOSHIDA, A. (1976) G6PD Lozere and Trinacria-like. *Hum. Genet.* **34**, 293-298.

WANNER, L. A., NEEL, J. V. and MEISLER, M. A. (1982) Separation of allelic variants by two dimensional polyacrylamide gel electrophoresis. *Amer. J. Hum. Genet.* **34**, 209-215.

WATKINS, W. M. (1966) Blood group subtances. *Science* **152**, 172-181.

WEATHERALL, D. J. (1978) The thalassemias. In: J. B. Stansbury, J. B. Wyngaarden and D. S. Fredrickson (eds.), *The Metabolic Basis of Hereditary Disease*, 4th ed. McGraw-Hill, New York.

WEATHERALL, D. J. and CLEGG, J. B. (1982) Thalassemia revisited. *Cell* **29**, 7-9.

WOMACK, J. E. and EICHEN, E. M. (1977) Liver-specific lysosomal acid phosphatase (Apl) ofmo use chromosome 17. *Mol. Gen. Genet.* **155**, 315-317.

WURZINGER, K. H. and MOHRENWEISER, H. W. (1982) Studies on the genetic and non-genetic (physiological) variation of human erythrocytes glutamic oxaloacetic transaminase. *Ann. Hum. Genet.* **46**, 191-200.

YOSHIDA, A. (1973) Hemolytic anemia and G6PD deficiency. *Science* **179**, 532-537.

YOSHIDA, A. (1977) Glucose-6-phosphate dehydrogenase abnormality and hemolysis. *Acta Biol. Med. Ger.* **36**, 689-701.

YOSHIDA, A. and BEUTLER, E. (1983) G6PD variants: another up-date. *Ann. Hum. Genet.* **47**, 25-38.

YOSHIDA, A., CHILLAR, R. and TAYLOR, J. C. (1979) An alpha-1-antitrypsin variant, Pi B Alhambra (Lys→Asp, Glu→Asp), with rapid anodal electrophoretic mobility. *Amer. J. Hum. Genet.* **31**, 555-563.

YOSHIDA, A., EWING, C., WESSELS, M., LIEBERMAN, J. and GAIDULIS, L. (1977) Molecular weight abnormality of PI S variant of human alpha-1-antitrysin. *Amer. J. Hum. Genet.* **29**, 223-239.

YOSHIDA, A., LIEBERMAN, J., GAIDULIS, L. and EWING, C. (1976) Molecular abnormality of human alpha-1-antitrypsin variant (Pi-ZZ) associated with plasma activity deficiency. *Proc. Nat. Acad. Sci. USA* **73**, 1324-1328.

YOSHIDA, A., TAYLOR, J. C. and VANDENBROEK, W. G. M. (1979) Structural differences between the normal Pi M_1 and the common Pi M_2 variant of the human alpha-1-antitrypsin. *Amer. J. Hum. Genet.* **31**, 564-568.

YOSHIDA, A., WATANABE, S., CHEN, S. H., GIBLETT, E. R. and MALCOLM, L. A. (1972) Human phosphoglycerate kinase. ii. Structure of a variant enzyme. *J. Biol. Chem.* **247**, 446-449.

YOSHIDA, A. and WESSELS, M. (1978) Origin of the multiple components of human alpha-1-antitrypsin. *Biochem. Genet.* **16**, 641-649.