Uses and misuses of definitions of genetic polymorphism. A perspective from population pharmacogenetics

Definitions of genetic polymorphism currently employed in the pharmacological literature have been incorrectly used (Kalow, 1984) and uncritically imported from *population genetics* [e.g., Weinshilboum (1984), Evans (1977)], apparently without regard to the problems their application might present in pharmacogenetics or, more specifically, in pharmacogenetic studies of populations.

Since a strong background in *population genetics* is not widespread amongst population pharmacogeneticists and especially amongst pharmacologists, there is a growing danger that these definitions will be misinterpreted or incorrectly used, as research on potentially polymorphic biotransformation routes increases.

The term polymorphism was first defined by E. B. Ford who stated that

> '. . . Polymorphism may be defined as the occurrence together in the same habitat of two or more discontinuous forms of a species in such proportions that the rarest of them cannot be maintained merely by recurrent mutation . . .' (Ford, 1940, 1965).

Ford, in fact, was referring to morphological characteristics, mostly those of butterflies (Lewontin, 1974).

The emphasis on phenotypic traits as a criterion for polymorphism is again reflected in the definition proposed by Vogel and Motulsky, who state:

> '. . . A polymorphism is a Mendelian or monogenic trait that exists in the population in at least two phenotypes (and presumably at least two genotypes), neither of which is rare—that is, neither of which occurs with a frequency of less than 1–2% . . . A polymorphism should be contrasted with a *rare* genetic variant. Rare genetic variants are arbitrarily defined as monogenic traits that occur in the population with a frequency of less than 1–2% and usually at much lower frequencies.' (Vogel & Motulsky, 1986).

The essence of this definition is contained in the first of two earlier ones proposed by Cavalli-Sforza & Bodmer (1971) (Chapter 2, p. 41). In a later chapter of their book, these authors present an alternate set of definitions, which has been adopted by others, including the frequently cited Harris (1980) and, more recently, Crow (1986) and Nei (1987). They defined polymorphism in the following way:

> 'Genetic polymorphism is the occurrence in the same population of two or more alleles at one locus, each with appreciable frequency.'

They go on to affirm that '. . . The definition of "appreciable frequency" is *arbitrary* (our italics) . . .', and then declare that '. . . it can be taken to be of the order of one percent.' (Cavalli-Sforza & Bodmer, 1971, Chapter 4, p. 118).

It is evident from these definitions that there is still no common agreement, but rather a very fierce controversy (Harris, 1980) about the evolutionary origin of polymorphisms and the forces which maintain them at specific frequency levels (Nei, 1987). As a consequence, all definitions of polymorphism used today are, as they ought to be, operative or working definitions since, starting from Cavalli-Sforza & Bodmer in 1971, they purposefully eschew any real theoretical underpinning. Hence, the differences between polymorphism and rare genetic variants are no longer explained by mechanistic arguments and are, therefore, necessarily and entirely arbitrary (Crow, 1986; Hartl, 1980; Hedrick, 1983; Nei, 1987).

According to the allelic definition of Harris (1980, p. 331), polymorphism exists when the '. . . commonest identifiable allele has a frequency no greater than 0.99 . . .' ($P \leq 0.99$ or $q \geq 0.01$). This signifies that, when the Hardy-Weinberg law applies ($p^2 + 2pq + q^2$) and when 1% is taken as the lower limit of frequency for the rarest allele, the distribution of phenotypes would result in 98.01% of homozygous individuals for the commonest allele, 1.98% of heterozygotes, and 0.01% (1 in 10,000 subjects) of homozygotes for the least frequent allele.

In the field of population pharmacogenetics, the main concern is the determination of phenotypes, because of the potential clinical implications arising from the large differences in metabolic activity between extensive and poor metabolizers, e.g., 20-fold in Caucasian sparteine oxidation. Furthermore, classical methodology (metabolic phenotyping) does not permit, in most cases, the resolution of homozygotes dominant (the so-called extensive metabolizers) from the heterozygous subjects (Tucker *et al.*, 1986) or the unambiguous identification of phenotypes with genotypes (Steiner *et al.*, 1985). This is valid even when the most up-to-date DNA probe studies are conducted (Idle, 1989). A further point in favour of the identification of phenotypes arises from the potential dissociation between phenotype and genotype, as might be the case in certain races (Yue *et al.*, 1989).

The desire to determine phenotypic status coupled with the common inability to detect genotypes directly, resulted in the universal adoption in the pharmacological literature of Cavalli-Sforza & Bodmer's (1971) and Vogel & Motulsky's (1986) phenotypic definitions for polymorphism. The only reasonable interpretation of these phenotypic definitions is that, if the number of poor metabolizers (presumably homozygotes recessive) corresponds to at least 1–2% of the total sample population, genetic polymorphism exists. Inversely, if the number of poor metabolizers is below 1–2%, polymorphism is said not to be evident or its absence could have been implied (Arias *et al.*, 1986; Eichelbaum &

Correspondence: Professor T. D. Arias, Apartado 10767, Estafeta Universitaria, Panamá, República de Panamá.

Woolhouse, 1985; Iyun *et al.*, 1986; Lou *et al.*, 1987; Nakamura *et al.*, 1985; Tucker *et al.*, 1986; Woolhouse *et al.*, 1985).

This view is shared by all researchers in population pharmacogenetics who have claimed absence of evidence for the existence of polymorphism, and, presumably, by the editors and most of the reviewers of their publications. Had pharmacogeneticists explicitly adopted the newer allelic-based definition, exclusively employed in the last 20 years in *population genetics*, they would have had to face the practical impossibility of studying thousands of subjects before they could claim that polymorphism was absent. Population geneticists have not encountered this limitation ever since efficient electrophoretic and, to a lesser extent, *ex vivo* techniques have been available. They can easily search both for the least common homozygous subjects, as well as for the almost 200 times more numerous heterozygotes. As a result, population geneticists have usually only needed to study '. . . no more than a hundred or two . . .' unrelated subjects to be able to detect polymorphisms (Harris, 1980, p. 343). This explains why this number has pervaded pharmacogeneticists' circles as an incorrect first approximation to the desired sample size, given the usual impossibility of identifying heterozygotes.

Pharmacogeneticists have failed to recognize that allelic and phenotypic definitions yield different estimates for phenotypes in a given population. It became evident to us—at least 3 years ago (Arias *et al.*, 1988)—that, when the Hardy-Weinberg law was used to calculate phenotypic frequencies at the lower limits of the allelic and phenotypic definitions of polymorphism, there was an unacceptable discrepancy. For instance, in the phenotypic definition the least common phenotype ($q^2 \geq 0.01$) would show a frequency one hundred times larger than the value ($q^2 \geq 0.0001$) which would result if Harris' allelic definition ($q \geq 0.01$) were applied. Similarly, there is an approximate 10-fold difference in heterozygote frequency (2pq).

While this paper was under review, we consulted with one of the authors of the most commonly used phenotypic definitions (A. Motulsky), who indicated the following (November, 1988):

> '. . . We (Vogel/Motulsky) may have contributed to the confusion by not expanding our definition of polymorphisms to indicate that we did not mean homozygotes when referring to a phenotype frequency of more than 1–2% . . . the Harris definition (q or allele frequency > 0.01–0.02) is similar to the Vogel-Motulsky definition which implies 2q or heterozygote frequency > 0.01–0.02 or q > 0.005–0.01 . . .'

Motulsky's overdue clarification makes it now necessary for pharmacogeneticists to abandon their phenotypic interpretation of Vogel & Motulsky's (1986) definition of polymorphism and search for an adequate definition.

For the reasons stated above, we are of the opinion that, in pharmacogenetics, definitions of polymorphisms must be explicitly based on phenotypic grounds, with a 1% frequency of the least common phenotype as the lower limit. This value is as arbitrary as any other ever used in *population genetics* (Cavalli-Sforza & Bodmer, 1971; Crow, 1986; Harris, 1980; Hartl, 1980; Hedrick, 1983; Nei, 1987); but it best lends itself to the interests and possibilities of our discipline.

In order to avoid contradictions between pharmacogenetic usage and genetic theory and concepts, we further propose that the polymorphisms thus characterized be referred to as 'pharmacogenetic polymorphisms'. The essential nature of these polymorphisms was recognized by Harris (1980, p. 340) when he described them as '. . . "quantitative" enzyme polymorphisms . . .'. They would be defined in terms of phenotypic frequencies, with a lower limit of 1%, and would be characteristic of metabolic inborn errors of pharmacologic relevance. This new definition would not require unambiguous identification of genotypes with phenotypes and some other Hardy-Weinberg conditions for its application, such as equilibrium. The latter feature would allow its use in populations which deviate from these characteristics, e.g., Amerindian groups and other genetic isolates.

Although the authors assume entire responsibility for the views stated in this letter, they wish, nevertheless, to thank the help received from Professors A. Motulsky, Doctors T. Inaba, R. G. Cooke and the anonymous reviewers of the manuscript. This work was supported by the Regional Program for Scientific and Technological Development, Organization of American States.

T. D. ARIAS[1], L. F. JORGE[1] & R. BARRANTES[2]
[1]The World Health Organization Collaborating Centre for Drug Quality Control, Instituto Especializado de Análisis, and the College of Pharmacy, University of Panamá, Panamá, Panamá and [2]Instituto de Investigaciones en Salud (INISA), University of Costa Rica, San José, Costa Rica.

Received 28 April 1989,
accepted 4 September 1990

References

Arias, T. D., Jorge, L. F., Lee, D., Barrantes, R. & Inaba, T. (1988). The oxidative metabolism of sparteine in the Cuna Amerindians of Panama: Absence of evidence for deficient metabolizers. *Clin. Pharmac. Ther.*, **43**, 456–465.

Arias, T. D., Jorge, L. F. & Inaba, T. (1986). No evidence for the presence of poor metabolizers of sparteine in an Amerindian group: The Cunas of Panama. *Br. J. clin. Pharmac.*, **21**, 547–548.

Cavalli-Sforza, L. L. & Bodmer, W. F. (1971). *The genetics of human populations*, pp 41 and 118. San Francisco: W. H. Freeman & Co.

Crow, J. F. (1986). *Basic concepts in population, quantitative,*

and evolutionary genetics, p. 14. New York: W. H. Freeman.

Eichelbaum, M. & Woolhouse, N. M. (1985). Inter-ethnic difference in sparteine oxidation among Ghanaians and Germans. *Eur. J. clin. Pharmac.*, **28**, 79–83.

Evans, D. A. P. (1977). Human pharmacogenetics. In *Drug metabolism—from microbe to man*, eds Parke, D. V. & Smith, R. L., pp 369–391. London: Taylor and Francis Ltd.

Ford, E. B. (1940). Polymorphism and taxonomy. In *The new systematics*, ed. Huxley, J., pp 493–513. Oxford: Clarendon Press.

Ford, E. B. (1965). *Genetic polymorphism*. Cambridge, MA: The MIT Press.

Harris, H. (1980). *Principles of human biochemical genetics*, 3rd edition, p. 331. New York: Elsevier/North Holland Biomedical.

Hartl, D. E. (1980). *Principles of population genetics*, p. 77. Sutherland, Mass.: Sinauer.

Hedrick, P. W. (1983). *Genetics of populations*, p. 67. Boston: Science Books International.

Idle, J. R. (1989). Poor metabolizers of debrisoquine reveal their true colours. *Lancet*, **ii**, 1097.

Iyun, A. O., Lennard, M. S., Tucker, G. T., Woods, H. F. & Phil, D. (1986). Metoprolol and debrisoquine metabolism in Nigerians: Lack of evidence for polymorphic oxidation. *Clin. Pharmac. Ther.*, **40**, 387–394.

Kalow, W. (1984). Pharmacoanthropology, outline, problems, and the nature of case histories. *Fed. Proc.*, **43**, 2314–2318.

Lewontin, R. C. (1974). *The genetic basis of evolutionary change*, p. 30. New York: Columbia University Press.

Lou, Y.-C., Ying, L., Bertilsson, L. & Sjoqvist, F. (1987). Low frequency of slow debrisoquine hydroxylation in a native Chinese population. *Lancet*, **ii**, 852–853.

Nakamura, K., Goto, F., Ray, W. A., & McAllister, C. B., Jacqz, E., Wilkinson, G. R. & Branch, R. A., (1985). Interethnic differences in genetic polymorphism of debrisoquine and mephenytoin hydroxylation between Japanese and Caucasian populations. *Clin. Pharmac. Ther.*, **38**, 402–408.

Nei, M. (1987). *Molecular evolutionary genetics*, pp 176–177, 405–431. New York: Columbia University Press.

Steiner, S. E. (1987). *Polymorphic debrisoquine hydroxylation. (Dissertation)*, pp. 9–10. Stockholm, Sweden: The Karolinska Institute.

Steiner, E., Iselius, L., Alvan, G., Lindsten, J. & Sjöqvist, F. A. (1985). A family study of genetic and environmental factors determining polymorphic hydroxylation of debrisoquine in man. *Clin. Pharmac. Ther.*, **38**, 394–401.

Tucker, G. T., Jackson, P. R., Lennard, M. S. & Woods, H. F. (1986). The detection of polymorphic drug oxidation—some theoretical and practical aspects. In *Interethnic differences in response to drugs and xenobiotics*, eds Kalow, W., Goedde, H. H., Agarwal, A. G., pp 413–424. New York: Alan R. Liss.

Vogel, F., Motulsky, A. G. (1986). *Human genetics. Problems and approaches*, 2nd edition, p. 435. New York: Springer Verlag.

Weinshilboum, R. M. (1984). Human pharmacogenetics. Introduction. *Fed. Proc.*, **43**, 2295–2296.

Woolhouse, N. M., Eichelbaum, M., Oates, N. S., Idle, J. R. & Smith, R. L. (1985). Dissociation of co-regulatory control of debrisoquine/phenformin and sparteine oxidation in Ghanaians. *Clin. Pharmac. Ther.*, **37**, 512–521.

Yue, Q. Y., Bertilsson, Dahl-Puustinen, M. L., Säwe, J., Sjöqvist, F., Johansson, I. & Ingelman-Sundberg, M. (1989). Disassociation between debrisoquine hydroxylation phenotype and genotype among Chinese. *Lancet*, **ii**, 870.

ADONIS 030652519100020P

Phenotype or genotype?

The term polymorphism means literally 'many shapes or forms'. In biology, polymorphism was used originally with respect to morphology but is now applied to any biological characteristic. In population studies polymorphisms are determined operationally by inspection of the frequency distribution of the trait in question. Any discontinuity in the distribution curve could legitimately be said to demonstrate the existence of a polymorphism. Although the underlying basis of such a polymorphism may be genetic, this is not necessarily so. Thus, strictly speaking, those polymorphisms with a genetic basis should be described as genetic polymorphisms, although the term polymorphism is now widely used in this sense.

Pharmacogeneticists may be divided into those who wish to use (genetic) polymorphisms as anthropological tools and those interested in their possible clinical relevance in interindividual variability in drug metabolism. The former group often wish to define differences in genotype between populations, whilst the latter are primarily interested in determining differences in phenotype between individuals. The difference between these aims is reflected in the way in which a polymorphism is described. Thus, the definition may be couched either in terms of the genotype, based on the frequency of the less common allele, or in terms of the phenotype, based on the frequency of the least common phenotype. This is not just a question of semantics, but has important implications for all areas of pharmacogenetics. In their letter, Professor Arias and his colleagues (1991) point out the confusion which has arisen from the indiscriminate use of these alternative, but not interchangeable, definitions, and argue in favour of the 'phenotype' definition largely because they consider it easier to implement. They indicate that the early definition of a (genetic) polymorphism, devised by population geneticists, was based on phenotype, but it is perhaps not surprising that this definition has 'evolved' as the science of genetics has progressed. Thus, when the only means of determining genotype was by breeding experiments, the phenotype was commonly used as the unit of genetic variation. Now, when it is possible to determine genotype directly, the allele has become the unit preferred by population geneticists. Whilst the mechanistic distinction between polymorphisms and rare genetic traits is unclear, the exact value of the frequency of occurrence of the least common genetic variant must remain arbitrary.