Theor. Appl. Genet. 56, 183-185 (1980)

Heterozygosity Levels in a Mammal: A Determination Derived from Allele Frequencies at Loci Controlling Gross Morphology

B. Blumenberg
Faculty of Sciences, Lesley College, Cambridge, Massachusetts (USA)

Summary. Estimates of heterozygosity derived from electrophoretic data may seriously underestimate the true levels of genetic variability. An estimate derived from allele frequencies in *Felis* that relate to macroscopic morphological traits appears to be congruent with several theoretical predictions of heterozygosity levels.

Key words: Allele frequency – *Felis* – Gross morphology – Heterozygosity

Introduction

During the last decade, it has become routine and established procedure to study genetic variation in natural populations by means of the technique of gel electrophoresis (Lewontin and Hubby 1966; Harris 1966). Such analysis reveals electrophoretic allozymic variation. In a recent and comprehensive review, Nevo (1978) surveyed the literature and computed a mean heterozygosity level (H) for mammals as 0.0359. The H levels in the tabulation range from 0.0 for *Mirounga anqustirostris* (and others) to 0.106 for *Mus musculus*.

To a minority of workers, a fundamental question remains. Are such determinations truly representative of the genome or do they merely reflect the limitations of available technique? Several authors, after considering both theoretical and practical aspects of such methodology, have aligned themselves with the latter view. King and Wilson (1975), Marshall and Brown (1975) and Nixon (1977) estimated that perhaps as few as 25% of all mutations are electrophoretically detectable. Bonhomme and Selander (1978) in a survey of variation in both electrophoretic charge and thermostability at 14 structural loci in *Mus musculus* estimated that electrophoresis detects about one-third of the variants which is equivalent to 30% of the alleles at such loci. An important factor certain to complicate this problem further stems from the observation that some isozymes show patterns of inheritance that cannot be interpreted by simple Mendelian models (Bowen and Yang 1978). Nei (1975: pp. 134-135) estimated that real heterozygosity levels in outbreeding organisms may be on the order of 30-40%.

Observations of H for individual loci suggest a wide range of values. Johnson (1974) provides a summary of such data. In small vertebrates, variable substrate enzymes range in H levels to 30% (esterases); regulatory enzymes to 24% (phosphoglucomutase 2,3) and non-regulatory enzymes to 13% (isocitrate dehydrogenase). In a recent review, Ward (1977) gives a mean vertebrate heterozygosity level for adenosine deaminase of 28%. In man, Johnson (1974) presents H levels for variable substrate enzymes that range upwards to 47% (pepsinogen), for regulatory enzymes that range to 50% (glutamate pyruvate aminotransferase) and for non-regulatory enzymes that range to 9% (amylase). Hedrick and Murray (1978) report H levels for 12 red blood cell loci in man that exceed 30%. Those loci showing heterozygosity levels in excess of 20% include representatives of both the fast and slow evolving groups (Sarich 1977). It is worth noting that a) loci producing regulatory enzymes do show lower levels of heterozygosity than all others and b) the proportion of any genome upon which heterozygosity (or indeed any other) determinations rest is well under 1%.

Methods

Is there any approach to heterozygosity conceptualization and calculation that results in mean H levels in excess of 20%? There appears to be one set of studies that provides data that lead to such a determination. Free ranging populations of the domestic cat (*F. catus*) have long been the subject of investigations in population genetics. These studies have focused upon an examination of several polymorphic loci the alleles of which determine gross morphological traits such as coat color, hair length, presence or absence of tail, manifestations of polydactyly, etc. A discussion of

0040-5752/80/0056/0183/$ 1.00

the specific alleles and their inheritance patterns may be found in Robinson (1977). Regional reviews may be found in Blumenberg (1976, 1977, 1978); Blumenberg and MacDonald (1978); Borodin et al. (1978); Morrill and Todd (1978); Todd and Blumenberg (1978); Todd and Kunz (1977) and Todd and Todd (1976a, b). All such alleles exhibit understandable patterns of Mendelian inheritance.

Due to the existence of an incompletely dominant sex-linked mutant (orange), which may be diagnosed in the heterozygous state, and the distinctiveness of the heterozygote at the piebald spotting locus (Dreux 1975), two tests for panmixia may be applied. Virtually all populations studied to date (in excess of 150) meet the requirements for panmixia (Hardy-Weinberg equilibria) and presumably adhere to the underlying structural and behavioral assumptions. No other mammal is known that a) possesses a variety of easily scored mutant alleles that result in gross morphological polymorphisms and b) also meets the requirements for panmixia in its breeding structure.

Allele frequency data was taken from Blumenberg and MacDonald (1978) and used as the basis for a gene diversity (heterozygosity) analysis according to Nei (1975: pp. 149-150). Nei (1975) defines such diversity as the frequency of heterozygotes under Hardy-Weinberg equilibrium and partitions the diversity into intrasubpopulational and intersubpopulational components. The analysis is designed to be applied to a large number of loci among a finite number of subpopulations and uses gene frequencies of the present generation. No assumptions are required about the pedigrees of individuals, selection or migration.

All that need concern us here is the case of two alleles (x, y) at each locus. Let x_{ik} be the frequency of the k^{th} allele in the i^{th} population. Then gene identity (1 – gene diversity) in this subpopulation is given by

$$J_i = \sum_k x^2_{ik} \quad (1)$$

and the gene identity in the total population is

$$J_T = \sum_k x^2_k \quad (2)$$

Gene diversity between the i^{th} and j^{th} populations may be defined as

$$D_{ij} = H_{ij} - (H_i + H_j)/2 = (J_i + J_j)/(2 - J_{ij}) = \sum_k (x_{ik} - x_{jk})^2/2 \quad (3)$$

where the gene diversity within the i^{th} population is give by

$$H_i = 1 - J_i \quad (4)$$

and the gene diversity between the i^{th} and j^{th} populations is given by

$$H_{ij} = 1 - J_{ij}$$

D_{ij} is identical to the minimum estimate of net codon differences between two populations. The total gene identity will reduce to

$$J_T = [\sum_i J_i]/s - [\sum_i \sum_j D_{ij}]/s^2 = J_S - D_{ST} \quad (5)$$

where J_S is the average gene identity within subpopulations and D_{ST} is the average gene identity between subpopulations including the comparisons of the subpopulations with themselves. H_o is the observed heterozygosity within colonies in the present generation

$$H_o = \sum_i \sum_j 2x_i y_i/s \quad (6)$$

All indices are averaged over all loci.

Results and Discussion

As seen in the example provided in Table 1, for those loci (*O*, *a*, t^b, *d*, *S*) that exhibit high levels of polymorphism (see regional reviews cited above), heterozygosity levels are quite high when compared to the results of many electrophoretic studies. Nearly identical values characterize these loci in each of the domestic cat populations that have been studied in the eastern half of the United States: Atlanta, Boston, Champaign, Chicago, Columbus (Ohio); Lawrence (Kansas), New York, Philadelphia, Portland, Rochester, St. Louis and Salem (Mass.) What is particularly striking is that this procedure yields *average* heterozygosity levels for each of these populations in excess of 32% for both interpopulational (H_i), intrapopulational (H_{ij}) and total (H_T) heterozygosity indices. One feature

Table 1. Gene diversity analysis

Comparison	locus*	J_i** ($\equiv J_S$)	J_j	J_{ij}	H_o	H_i** ($\equiv H_S$)	H_j	H_{ij}	D_{ij}** ($\equiv D_{ST}$)	J_T	H_T	$\bar{H}_S/\bar{H}_T$
Cleveland	*O*	0.64	0.59	0.61	0.39	0.36	0.41	0.39	0.0034	0.64	0.36	
	a	0.62	0.66	0.64	0.36	0.38	0.34	0.36	0.0018	0.62	0.38	
Portland Maine	t^b	0.53	0.50	0.51	0.49	0.47	0.50	0.49	0.0061	0.52	0.48	
	d	0.52	0.50	0.50	0.49	0.48	0.50	0.50	0.0042	0.51	0.49	
	S	0.54	0.51	0.52	0.47	0.46	0.49	0.48	0.0074	0.54	0.46	
	W	0.96	0.94	0.95	0.05	0.04	0.06	0.05	0.0001	0.96	0.04	
	Pd	0.96	0.94	0.95	0.05	0.04	0.06	0.05	0.0001	0.96	0.04	
	$\bar{X}$	0.68	0.66	0.67	0.33	0.32	0.34	0.33	0.0033	0.67	0.32	0.990

* The mutants considered are sex-linked orange (*O*), nonagouti (*a*), blotched tabby (t^b), blue dilution (*d*), piebald spotting (*S*), dominant white (*W*) and polydactyly (*Pd*); see Robinson (1977)

** Number of subpopulations has been set equal to 1

of this analysis that would introduce a bias towards high H levels is the absence of monomorphic loci, the inclusion of which commonly enters into such calculations when they are based upon electrophoretic data.

The procedure employed here violates two 'rules of the game' as established by workers proceeding from a foundation based in the study of allozymic variation.

1) The number of loci defining the data base is quite low. Only five to seven alleles entered into the calculations for each population.

2) Such allele frequencies do not represent cistrons coding for monomeric proteins or subunits of multimeric proteins. As pointed out in Todd and Blumenberg (1978), each allele frequency may represent an average frequency appropriate to each cistron that is producing enzymes within a particular biochemical pathway(s).

Searle (1968: pp. 49-52) recognizes three steps in the development of the melanin granule within the melanocyte involving the participation of six structures, only one of which is the protein producing polysome. Many, if not all, of the stages that comprise the ontogeny of the melanin granule must require the involvement of genetic activity, as observations of mutant allele effects upon melanocyte structure suggest (Searle 1968). Allele frequencies as determined for the *F. catus* genome by observations of gross morphological plymorphisms, may represent an average frequency for each of many active genes, both those controlling steps in the pathways leading to eumelanin and phaeomelanin synthesis and those involved in the subsequent formation of the pigment granule within the melanocyte. Searle (1968: p. 65) has provided a preliminary statement of the stages in the ontogeny of hair pigmentation at which particular coat color genes act in the mouse. Some of these loci are homologous to those in *Felis* (Searle 1975).

Little correspondence may be expected to exist between allele frequencies derived from the study of gross morphological polymorphisms and those obtained from electrophoretic protein studies. Nonetheless, the former data lead to estimates of heterozygosity in *Felis* that seem congruent with predicted levels of heterozygosity as deduced from both theoretical considerations and reflections upon the limitations of electrophoretic technique.

Literature

Blumenberg, B.: Mutant allele frequencies in the domestic cats of California, Texas and New Spain. Carniv. Genet. Newsl. **3**, 2-9 (1976)

Blumenberg, B.: Genetic difference and selection in domestic cat populations of the United Kingdom and former British colonies. Theor. Appl. Genet. **49**, 243-247 (1977)

Blumenberg, B.: Mutant allele frequencies in the domestic cats of eastern Massachusetts. Genetica **49**, 9-14 (1978)

Blumenberg, B.; MacDonald, B.: A population genetics analysis of the domestic cats of Cleveland, Ohio and the Midwest United States, Proc. First Int. Conf. Dom. Cat Pop. Genet. Ecol. Carniv. Genet. Newsl. **3**, 227-240 (1978)

Bonhomme, F.; Selander, R.K.: Estimating total genic diversity in the house mouse. Biochem. Genet. **16**, 287-297 (1978)

Borodin, P.M.; Bochkarev, M.N.; Smirnova, I.S.; Manchenko, G.P.: Mutant allele frequencies in domestic cat populations of six Soviet cities. J. Hered. **69**, 169-174 (1978)

Bowen, B.S.; Yang, S.Y.: Genetic control of enzyme polymorphisms in the California Vole, *Microtus californicus*. Biochem. Genet. **16**, 455-467 (1978)

Dreux, Ph.: Génétique de population des chats domestiques de Marseille (Bouches-du-Rhône, France). Ann. Génét. Sél. Anim. **7**, 23-33 (1975)

Harris, H.: Enzyme polymorphisms in man. Proc. Roy. Soc. (Lond.) B **164**, 298-310 (1966)

Hedrick, P.W.; Murray, E.: Average heterozygosity revisited. Am. J. Hum. Genet. **30**, 377-382 (1978)

Johnson, G.B.: Enzyme polymorphism and metabolism. Science **184**, 28-37 (1974)

King, M.C.; Wilson, A.C.: Evolution at two levels in humans and chimpanzees. Science **188**, 107-116 (1975)

Lewontin, R.C.; Hubby, J.L.: A molecular approach to the study of genic heterozygosity in natural populations. II. Amount of variation and degree of heterozygosity in natuural populations of *Drosophila pseudobscura*. Genetics **54**, 595-609 (1966)

Marshall, D.R.; Brown, A.H.D.: The charge-state model of protein polymorphism in natural populations. J. Mol. Evol. **6**, 149-163 (1975)

Morrill, R.B.; Todd, N.B.: Mutant allele frequencies in domestic cats of Denver, Colorado. J. Hered. **69**, 131-134 (1978)

Nei, M.: Molecular population genetics. New York: Elsevier 1975

Nevo, E.: Genetic variation in natural populations: patterns and theory. Theor. Pop. Biol. **13**, 121-172 (1978)

Nixon, S.E.: On the number of mutations that are detectable by electrophoresis: transition vs transversion mutations. Wasmann J. Biol. **35**, 77-86 (1977)

Robinson, R.: Genetics for cat breeders. 2nd ed. New York: Pergamon Press 1977

Sarich, V.M.: Rates, sample sizes and the neutrality hypothesis for electrophoresis in evolutionary studies. Nature **265**, 24-28 (1977)

Searle, A.G.: Comparative genetics of coat color in mammals. New York: Academic Press 1968

Searle, A.G.: Distribution and probable homologies of some genes determining coat color in mammals. In: Handbook of Genetics. (ed. King, R.C.) pp. 459-467. New York: Plenum Press 1975

Todd, N.B.; Blumenberg, B.: Mutant allele frequencies and genetic distance relationships in domestic cat populations of lower Egypt and the eastern Mediterranean. Theor. Appl. Genet. **52**, 257-262 (1978)

Todd, N.B.; Kunz, T.H.: Mutant allele frequencies in domestic cat populations of Greece. Biol. Gallo-Hellinica **6**, 289-310 (1977)

Todd, N.B.; Todd, L.M.: Mutant allele frequencies among domestic cats in some eastern areas of Canada. J. Hered. **67**, 368-372 (1976a)

Todd, N.B.; Todd, L.M.: Mutant allele frequencies in the domestic cats of Turkey and Greece. Genetica **46**, 183-192 (1976b)

Ward, R.D.: Relationship between enzyme heterozygosity and quaternary structure. Biochem. Genet. **15**, 123-135 (1977)

Received June 12, 1979
Communicated by D. Van Vleck

Prof. Dr. B. Blumenberg
Faculty of Sciences
Lesley College
29 Everett Street
Cambridge, Mass. 02238 (USA)