

personal communication). The careful monitoring of Etna's output is clearly important in interpreting these chemical observations in terms of the volcano's internal plumbing system.

With regard to the prediction of an eruption in May 1982  $\pm 3$  months, the *intermittent* presence of lava at the bottom of the Bocca Nuova (20 May–3 June, 9–20 August, 25–30 (?) September, observations by G. Patanè, Istituto di Scienze della Terra di Catania, and by members of the PIRPSEV teams) must be related to the normal oscillatory motion of the magma column. This activity has not brought any appreciable amount of new material as yet (5 December 1982). Similar activities occurred in 1981 (mid-May, mid-July, early September) and during summer 1980, including the SE Crater, (refs 2, 3): they were not interpreted as 'eruptions' by scientists and people concerned with Mount Etna. (At any rate, the prediction of an eruption with  $\pm 3$  months uncertainty is not significant for a volcano which erupts almost each year and often several times a year. Even had there been an eruption within this 6 months interval, it could have been a mere coincidence.)

G. WADGE  
J. E. GUEST

University of London Observatory,  
Mill Hill Park, London NW7 2QS, UK

1. Romano, R. & Sturiale, C. *Phil. Trans. B. Soc. A274*, 37–43 (1973).
2. *SEAN Bull. 7/5*, 12 (Smithsonian Institution, 1982).
3. *SEAN Bull. 7/7*, 8 (Smithsonian Institution, 1982).

## Identity of different mutations for deleterious genes

IN an analysis of  $\beta$ -thalassaemia in Mediterranean populations, Orkin *et al.*<sup>1</sup> have suggested that only eight separate mutational events leading to gene dysfunction are represented in 91 chromosomes taken from  $\beta$ -thalassaemic patients. This is perhaps the first time that such a statement on individual mutations has been possible on samples from human populations. They show further that there is a close association between mutational events and the pattern of restriction enzyme cleavage sites in the neighbourhood of the  $\beta$ -globin gene cluster.

We present here a theoretical treatment of this problem of identity of deleterious mutations in the framework of the classical 'infinite alleles' model of variation in finite populations, due to Kimura and Crow<sup>2</sup>. They assume that at the locus concerned each mutation is to a novel type and that all genotypes have equal fitness. If  $u$  is the total mutation rate per generation and  $N$  is the effective population size, they show that at equilibrium the probability that two genes chosen at random are identical by descent is  $1/(4Nu + 1)$ . This expression can be simply derived by an equation in which

the reduction in identity by new mutation is balanced against the increase due to finite population size. If  $f_t$  is the probability that a randomly chosen pair of genes in generation  $t$  are identical by descent, we may write

$$f_{t+1} = (1-u)^2(1/2N + (1-1/2N)f_t) \quad (1)$$

where  $(1-u)^2$  is the probability that neither is a new mutation and  $1/2N$  is the probability that they both derive from the same gene in generation  $t$ . Putting  $f_{t+1} = f_t$  leads to  $f = 1/(4Nu + 1)$ .

Consider a group of recessive lethal mutations at a locus which do not complement one another and which have equal fitness in heterozygotes with wild-type alleles. If  $u$  is the total mutation rate to such alleles and  $q$  is their total frequency, then the probability that neither gene of a pair is a new mutation is  $(q/(q+u))^2 = (1-u/q)^2$ . The number of mutant alleles in the population is  $2Nq$ , so that the expression  $1/2N$  in (1) must be replaced by  $1/2Nq$ . Thus the probability of identity at equilibrium is again  $1/(4Nu + 1)$ , with  $u$  now the mutation rate to lethal alleles. Note that this is independent of any selection on the heterozygote over wild type (provided this is the same for all alleles) as  $q$  can be taken as the frequency of mutant alleles after selection. The approach does in fact apply to any homogeneous subset of mutants at a locus.

The probability of identity will be increased by non-random mating, for example in local populations with limited migration, and in an expanding population it will be higher than current population sizes would suggest. The probability is also changed if not all deleterious homozygous types are equally unfit or there is any complementation of different mutant alleles.

For lethal genes, the expected value of  $q^2$  equals the total mutation rate,  $u$ , giving  $f = 1/(4Nq^2 + 1)$  or  $f = 1/(4L + 1)$ , where  $L$  is the total number of affected individuals (called clinical homozygotes by Orkin *et al.*<sup>1</sup>) in the population. The expected number of affected individuals which are homozygous for the same mutation is

$$Lf = L/(4L + 1) = Nu/(4Nu + 1)$$

which is less than one-quarter for all  $Nu$ . Thus, with lethality, the number of identical homozygotes is expected to be very small. If there is some heterozygote superiority for the mutant types, the number of affected individuals exceeds  $Nu$  and the number of identical individuals may exceed one-quarter.

Over a 60-kilobase length spanning the  $\beta$ -globin gene cluster, Orkin *et al.*<sup>1</sup> found in their sample of 91 chromosomes carrying  $\beta$ -thalassaemia mutations only nine patterns of presence or absence at a set of seven polymorphic restriction enzyme recognition sites, which they called 'chromosomal haplotypes'. The most

frequent haplotype made up 47% of the total. By sequencing the DNA of the  $\beta$ -thalassaemia genes on a number of chromosomes, they found only eight different types of mutations and, further, there was almost a one-to-one association of haplotype and thalassaemia mutation. As the frequency of haplotypes is similar in a sample of chromosomes with normal  $\beta$ -globin genes<sup>3</sup>, this implies that the thalassaemic mutations have occurred in a background of previously existing haplotype variation.

If this is so, we believe that they may have underestimated the number of thalassaemic mutational events by their experimental strategy as they deliberately concentrated their sequencing work on new haplotypes and only sequenced four samples of the most frequent haplotype (their Table 3): "Repeated isolation of the same gene was largely avoided by selecting  $\beta$ -genes for sequencing on the basis of their associated haplotypes". The actual probability of identity of mutants is likely to be lower than the value of 0.26, the observed identity of the restriction haplotypes. It is surprising that the frequency of haplotypes should be similar in normal chromosomes and in those carrying thalassaemia mutants and yet the more frequent haplotypes do not have proportionally more types of mutation.

ALAN ROBERTSON  
WILLIAM G. HILL

Institute of Animal Genetics,  
University of Edinburgh,  
Edinburgh EH9 3JN, UK

1. Orkin, S. H. *et al. Nature* **296**, 627–631 (1982).
2. Kimura, M. & Crow, J. F. *Genetics* **49**, 725–738 (1964).
3. Antonarakis, S. E., Boehm, C. D., Giardina, P. J. V. & Kazazian, H. H. Jr *Proc. natn. Acad. Sci. U.S.A.* **79**, 137–141 (1982).

ORKIN *ET AL.* REPLY—Robertson and Hill show mathematically that if there is an infinite number of lethal, non-complementary alleles at a locus such as  $\beta$ -globin, then the expected number of identical homozygotes in the population is less than 0.25. Thus, in the presence of random genetic drift, irrespective of the population size one will rarely find identical homozygotes and only occasionally more than one. In other words, under strict lethality nearly all affected individuals will be genetic compounds. This, of course, will not be true if there is some inbreeding in the population.

These authors argue that sequencing of  $\beta^{\text{thal}}$  alleles from the same haplotype will yield new and different  $\beta^{\text{thal}}$  alleles. In Orkin *et al.* haplotype I accounts for about 50% of the haplotypes associated with  $\beta^{\text{thal}}$  in Mediterraneans. Robertson and Hill suggest that this haplotype, because it is frequent, might be expected to carry several different  $\beta^{\text{thal}}$  alleles. We now have definitive information regarding  $\beta^{\text{thal}}$  alleles in 13 of these haplotype I chromosomes. Four  $\beta^{\text{thal}}$  genes from this haplotype were sequenced and all had the