Mixing Up the Genome

from our Molecular Genetics Correspondent

EVER since Britten and Kohne showed in 1968 that eukaryotic DNA consists both of sequences present only once and also of those which are repeated a large number of times in related but not identical copies, one of the most important problems in biology has been to define the function of the two types of component. One parameter which must first be defined is the arrangement of repeated and non-repeated sequences in DNA. Davidson, Hough, Amenson and Britten now show (*J. mol. Biol.*, **76**, 1; 1973) that the two types of component are interspersed in the genome of *Xenopus* and are able for the first time to make some estimate of the average length of each class of sequence.

When renatured in solution, the DNA of *Xenopus* displays two principal components. About 55% of the genome consists of non-repeated sequences which can renature only at high C_0t values produced by incubation at a large product of DNA concentration and time in solution. The remaining 45% of the DNA renatures at lower C_0t values, most of this fraction displaying an average degree of repetition of 1,600 copies, although two smaller fractions have smaller and larger degrees of repetition, respectively.

Renaturation experiments of this nature are usually made by using a standard preparation in which DNA is fragmented to a specified length, which has been 450 nucleotides in characterization of the *Xenopus* genome. Davidson *et al.* have now gained additional information by varying this technique; they have fragmented radioactively labelled preparations of DNA to different lengths (3,700, 1,400 or 700 nucleotides) and renatured them with an excess of unlabelled standard DNA of length 450 bases. The amount of radioactive duplex formed—measured by its retention on columns of hydroxyapatite—estimates the extent of reaction of the long fragments with the standard 450 length molecules.

By using low C_0t values, only repeated sequences of DNA are allowed to renature. But although only 45% of the genome consists of repeated sequences, 80% of the fragments of 3,700 base length were retained on the columns. These fragments, in addition to containing the repeated sequences which are responsible for the hybridization reaction, must therefore also include a large proportion of the non-repeated sequences. This conclusion was confirmed by eluting the bound 3,700 fragments from hydroxyapatite, shearing them to the standard 450 base length, and then rehybridizing. Only 46% of them continued to bind. The fragmentation must therefore have released some of the non-repeated sequences from the repeated sequences.

The quality of base pairing—a characteristic which defines the relationship between the repeated sequences which suffer hybridization—can be estimated by the melting temperature of renatured DNA. According to this criterion, renatured fragments of 1,400 bases and 450 bases long respectively implicate reaction of the same repeated sequences, for they have the same melting temperature. But the change in optical density upon melting the duplex—a parameter which identifies the extent of duplex formation—is twice as great with the fragments of 450 bases long. The fragments of 1,400 bases long must therefore form duplex DNA for a much lower proportion of their length. This implies that the increase in length from 450 to 1,400 bases allows the inclusion of greater amounts of non-repeated sequences which cannot renature under conditions of low C_0t. Taking into account both the extent of mispairing revealed by the melting temperature and the length of duplex formation revealed by the change in optical density, Davidson *et al.* calculate that the average length of the repeating unit which hybridizes is 300 ± 100 bases.

When long fragments (in these experiments 3,700 bases) are annealed with an excess of short fragments (the standard 450 base preparation), the rate of reaction of the radioactive preparation depends on the number of sites in each long fragment to which short molecules can bind. Because under appropriate conditions the 3,700 fragments react with 450 fragments at the same rate at which the 450 fragments react with each other, there seems to be no clustering of repeated sequences. There can, therefore, be no more than one or two repeated sequences in each 3,700 base long fragment.

By analysing the dependence of the proportion of fragments with repeated sequences upon the length of the preparation, Davidson *et al.* estimate that short repeated sequences of average length 200–400 bases alternate with longer non-repeated elements of average length 650–900 nucleotides. They suggest that there may be a high degree of order in the alternation of the two components. Experiments with other organisms in which repeated sequences have been characterized by electron microscopy of preparations annealed at low C_0t values suggest similar short lengths for this component; the range of lengths of the non-repeated sequences probably include those of the size of single genes, but would not extend to clusters of genes.

An obvious speculation upon the significance of these results is that the short repeated sequences represent the control elements which regulate the action of the adjacent non-repeated sequences. Because in *Xenopus* and in other organisms also much of the repeated component seems to be transcribed, its action cannot be confined to acting as elements for recognition by regulator molecules. But the alternating arrangement of repeated and non-repeated components is consistent with the concept that individual units of gene transcription may include both types of component; although the non-repeated sequence presumably codes for protein, the function of the repeated sequence must at present remain a matter for speculation.

CHROMOSOMES

Human Genes on the Map

by our Special Correspondent

ANYBODY who has ever wondered whether somatic cell hybridization would vindicate the enthusiasm of its early protagonists should note that this technique has so far made possible the assignment of about forty genes to various human chromosomes. Not all of these assignments are confirmed, however; some are only provisional, as was made clear at the first international workshop on human gene mapping at Yale University on June 10–13, under the auspices of the National Foundation—March of Dimes.

Although the cell hybridizers were much in evidence at the meeting, they did not manage to submerge the contributions of the geneticists who are continuing the valuable, if perhaps less fashionable, practice of mapping genes on the basis of studies of family pedigrees. In the true spirit of a workshop, much of the three days consisted of informal committee sessions, with two interludes for the communication of short papers.

A committee which spent half a day reviewing the genetic constitution of chromosome no. 1 agreed that it carries ten confirmed genetic loci: their criteria required that confirmation be based on studies by more than one laboratory on more than one family or by more than one method. (The methods of cell hybridization have been described by the workshop's organizer, Dr F. H. Ruddle, in *Nature*, **247**, 165; 1973.) Speaking for the committee, Dr Hamerton said they decided that six assignments to this chromosome must remain provisional. After plenary discussion one of them, the locus for adenylate kinase-2 (AK_2), was removed to its own