

In contrast to other positive strand RNA viruses, the corona-viruses readily establish persistent infections in the host. For example, chronic hepatitis or chronic demyelination of the central nervous system occur with certain virus-host combinations. Not surprisingly, persistent infections are relatively easy to establish in cell culture *in vitro*, and such model systems have some unusual features. K. Holmes (Bethesda), N. Hirano (Morioka) and S. Stohman (Los Angeles) each reported that in cloned cell cultures persistently infected with MHV only 10–20 percent of the cells express coronavirus antigens detectable by immunofluorescence, but 100 percent are resistant to superinfection with homologous virus. G. Chaloner-Larsson (Ottawa) has obtained very similar results with a human

coronavirus (229E) persistent infection of L 132 cells. Further characterisation of such *in vitro* systems is sure to provide much-needed insight into the way these viruses establish such an intimate relationship with the host. Summing up the meeting, D. Tyrrell (Harrow) emphasised the very wide range of diseases now known to be caused by these viruses, and the value of their study as models of virus pathogenesis. As the meeting convener, V. ter Meulen (Würzburg) pointed out, virtually every laboratory currently working with coronaviruses was represented amongst the seventy participants, and the collective presentation of their data undoubtedly served to strengthen and unify their purpose in unraveling the molecular biology of this unique group of viruses.

transposition is reasonably low, it should be possible to build a *Drosophila* strain that lacks copia sequences. Young has made a start to this interesting, but very tedious experiment: at the time of the meeting a strain with only six copia sequences had been made and the flies were apparently no worse for their experience. Only time, and hard work, will tell whether or not copia sequences are necessary for some aspect of *Drosophila's* existence.

The second genetic approach stems from the exciting discovery that at least one well known mutation (white-apricot) is associated with an insertion of copia into the white gene. B. Judd and P. Bingham (Research Triangle) have mapped this copia element within white to the same region as maps the  $w^{al}$  allele and W. Gehring's laboratory (Basel) has found that revertants of  $w^{al}$  may be accompanied by loss of the copia element from this site.

The role of the insertion of nomadic middle repeat sequences into genes as a cause of spontaneous mutation in *Drosophila* was also highlighted by the molecular study of the bithorax region. W. Bender (Caltech) has now cloned over 145 kilobase pairs of this region and has used these clones to analyse the nature of some of the many mutant bithorax genes. Two ( $bxd^1$  and  $bxd^5$ ), both spontaneous in origin, result from large insertions (of about 10 kb) into the gene. Although the insertion sites are not identical they are reasonably close together and each insertion is a repetitive DNA sequence. A spontaneous reversion of  $bxd$  turns out to have lost most, though not all, of its insertion. It must be most gratifying to E. B. Lewis, who is responsible for the very detailed genetic analysis of this complex locus, to see how well the molecular analysis of these mutations is confirming the genetic data. For example, the dominant mutation *Cbx* was recovered, after X-ray mutagenesis, together with a recessive *pbx* allele: Bender has now shown that *pbx* is a deletion and that the sequence removed from the *pbx* site has been inserted some kilobase pairs away in reverse orientation to give the *Cbx* mutation. Reversion of *Cbx* has been accompanied by an inversion broken within this inserted sequence.

It is to be hoped that the molecular analysis of long regions of the genome will throw some light on the old question of the relationship between genetic organisation and chromosome structure. P. Spierrer (Geneva) recounted his latest travels along the chromosome in the region of the genes coding for xanthine dehydrogenase (*rosy*) and acetyl cholinesterase. Over 300 kb of DNA (three 'Benders') have been covered and this spans eight polytene chromosome bands and as many characterised genes. As dramatically as any other this experiment demonstrates the variation in location of middle repeat DNA sequences. The DNA cloned for this walk came from two different 'wild type' stocks of *D.*

## Drosophila at Kolymbari

from M. Ashburner

FOR the second time the Greek Orthodox Academy at Kolymbari, on the northern shores of Crete, was host to a meeting of *Drosophila* biologists\*. Not surprisingly, the scientific sessions were dominated by discussions of the molecular structure of the *Drosophila* genome. Restriction maps and nucleotide sequences were so common that I am sure more than one participant considered that his colleagues, but naturally not himself, had fallen into the trap so carefully avoided by Dr Watson, of confusing hard work with hard thinking. Nevertheless some very interesting facts were brought to our attention: interesting if only for the reason that they were quite unexpected.

There can be little doubt that the major challenge to conventional genetic wisdom revealed by the molecular analysis of the genome of *Drosophila*, and other eukaryotes, has been the discovery of repeated DNA sequences of no fixed abode. The 'type' of these sequences, as far as *Drosophila* is concerned, is that dubbed 'copia', a sequence of some 5 kilobase pairs which is present, in the typical strain of *D. melanogaster*, roughly 30 times. Comparison of the chromosomal locations of this sequence in different strains of *D. melanogaster* has shown that these differ between, and even within, strains.

Studies on the sequence organisation of copia, and similar elements, by G. Rubin's laboratory (Harvard Medical School) have revealed striking similarities between copia, a dispersed middle repeat sequence of yeast known as Tyl, and integrated vertebrate retroviruses. Copia is flanked by

a direct terminal repeat of 276 base pairs and a comparison of the nucleotide sequence of sites into which copia has inserted reveal that insertion is accompanied by a direct duplication of five base pairs of the 'target' sequence. Other nomadic middle repeat sequences of *Drosophila* appear to have a similar overall organisation to copia but to generate duplications of the insertion site of different sizes (for example four or seven base pairs). The similarities between these elements and bacterial IS sequences are striking; whether these similarities reflect the mechanisms IS sequences and the nomadic eukaryotic elements use for transposition is not yet known.

Copia, and some other nomadic middle repeat sequences of *Drosophila*, were originally cloned by virtue of the fact that their RNA transcripts are particularly abundant, at least in tissue culture cells. Work in both M. Young's laboratory (Rockefeller University) and in Rubin's has now shown that several different transcripts complementary to copia can be found in cells: whether these can all come from any single copia sequence is not known. For many years the function of RNAs complementary to copia has been the subject of some considerable discussion. They are clearly not translated in proportion to their mass yet, as work by both Rubin and J. Lengyel (University of California, Los Angeles) has shown, these RNAs can be translated into polypeptides in a cell free system.

Two genetic approaches to understanding the function of these nomadic sequences were discussed at Kolymbari. Since different stocks of *D. melanogaster* differ in the location of their copia sequences, and since the frequency of copia

M. Ashburner is in the Department of Genetics, University of Cambridge.

*melanogaster*: Oregon-R and Canton-S. Overall seven different insertions of middle repeat DNA were found but none were in the same location in both stocks: indeed many were polymorphic within a stock. Yet for non-repetitive DNA these two stocks show only a fraction of a percent divergence in sequence.

A rather different type of transposable element in *Drosophila* was described by G. Ising (Lund). This element, called TE1, is huge, it carries at least two genes (white and roughest) normally adjacent in the X. It arose spontaneously some years ago as a transposition of  $w^a$  and  $rst^+$  from the X chromosome to chromosome two and, since then, has been merrily hopping around the genome with a frequency of about  $1:10^4$ . Ising has mapped over 150 different sites into which it has inserted, insertion often being accompanied by mutation of the 'target' gene. Just how TE1 moves, or how it can be lost altogether from the genome, is rather mysterious but there is hope that the cloning of one end of TE1 fortuitously inserted very near one of the heat shock genes (M. Goldberg, Basel) will help to solve this mystery.

As at Kolymbari two years ago the analysis of genes of known function, known at least in terms of their protein products, continues to give great pleasure to participants and onlookers alike. Those few genes in *Drosophila* whose activity is so dramatically induced by heat shock are by far the best known. Cloning has confirmed earlier genetic conclusions that those coding for the four small heat shock proteins are very closely linked. R. Morimoto (Harvard) described that their coding sequences are all contained within a 13 kb fragment. Detailed comparison of

the DNA sequences upstream of the 5' ends of these genes with those of genes coding for three other heat shock proteins is revealing homologies some 125 bases upstream from the 'TATA box' sequence.

The clones of the 'major' heat shock gene (Hsp70), a gene that is repeated at each of two different loci in *D. melanogaster*, are now beginning to be used for a number of quite different studies, for example of gene evolution (D. Ish-Horowicz, ICRF, London), nucleosome spacing (M. Noll, Basel) and chromatin organisation revealed in the pattern of sensitivity to nucleases (S. Elgin and C. Wu, Harvard). Yet early hopes that a comparison of the sequences adjacent to different Hsp70 copies would quickly reveal those sequences important for the control of transcription appear to have been too optimistic, since variant Hsp70 genes have been isolated that can be shown to be functional, yet lack a characteristic class of sequences normally found 5' to the coding region (M. E. Mirault, Geneva). On the other hand the conference heard from J. Liss (Cornell) that the insertion of a small (406 base pair) sequence, which includes the first 63 base pairs of the translated Hsp70 gene and material 5' to it, into a particular repetitive DNA sequence found adjacent to some Hsp70 genes results in the induction, by heat shock, of transcription of the middle repeat sequence.

Unfortunately, the function of the heat shock proteins remains unknown. Their importance for the cell in the recovery of normal protein synthesis after the heat shock was shown by S. Lindquist (Chicago). The incorporation of certain amino acid analogues into the heat shock

proteins apparently blocks their post-transcriptional processing and normal accumulation in the nucleus; it also blocks the recovery of normal protein synthesis after the heat shock. Lindquist suggests that normally these proteins may be nucleic acid binding proteins. Recently, rather similar responses to environmental shock have been found in organisms as different as yeast and man. A. Tissières (Geneva) discussed evidence from fingerprints of 'heat shock' proteins from different animals that indicates that at least Hsp70 may be conserved.

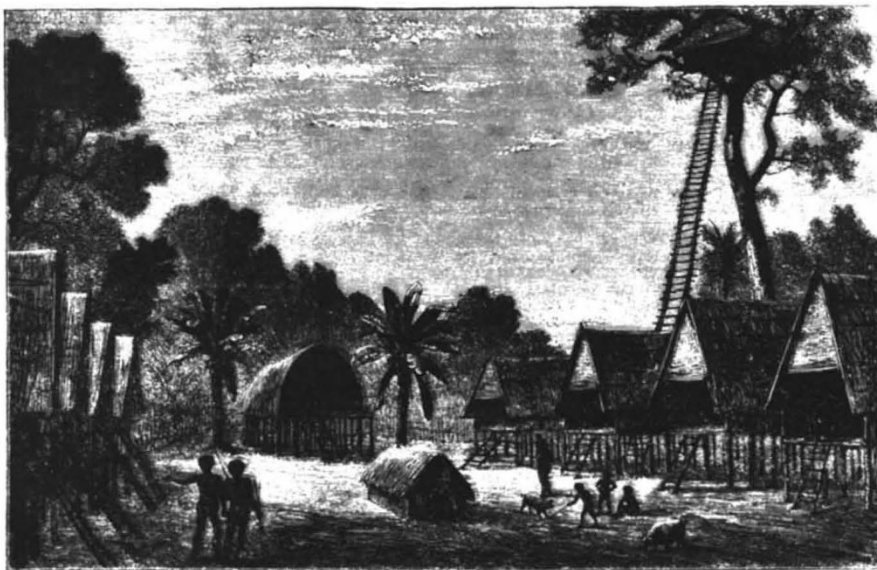
Progress is also slow in discovering the mechanisms of induction of heat shock protein synthesis. T. Kornberg (University of California, San Francisco) has characterised an *in vitro* system using isolated *Drosophila* tissue culture cell nuclei in which the *E. coli* RNA polymerase directed transcription of the heat shock genes is responsive to the addition of factors from heat shocked, but not control, cells. The nature of these factors (or factor) is yet to be elucidated.

Other genes that attracted attention at Kolymbari were those coding for the salivary gland 'glue' proteins and the proteins of the egg shell of *Drosophila*. Each had their surprises. A component of the salivary gland glue, known as SGS-4, is coded for by a gene on the X chromosome and is of interest for several reasons. G. Korge (Munich) has characterised mutations that normally fail to express this gene yet will do so when the mutant gene is made heterozygous with a wild type allele, at least under certain conditions. This is rather similar to the transvection affect described for the bithorax locus by E. B. Lewis 25 years ago. The coding region for SGS-4 has

100 years ago



Fanduri — a Dorey papuan



Epa — a Village of the Mahori papuans.

Pictures from "New Guinea: What I did and what I saw" by L. M. D'Alberis reported in *Nature* 23, 16 December, 153, 1880.

been cloned in Stanford and is being used for the analysis of this entire region of the genome by S. Beckendorf (University of California, Berkeley) and for detailed studies of the gene itself by M. Muskavitch (Stanford). A surprising result is that the transcript of SGS-4 differs in size in different strains of *Drosophila*. This results, it would appear, from the fact that within the coding sequence there is a 21 base pair sequence whose repetition frequency can vary between different SGS-4 alleles. Although nothing is known of the chemistry of the protein this must reflect a sequence of seven amino acids repeated a varying number of times.

The mutation characterised by Korge, that fails to make SGS-4 protein when homozygous turns out to be a small deletion, but of a sequence over 200 base pairs away from the 5' end of the messenger RNA. The importance of this region is indicated by the fact that two different expressed alleles of SGS-4 differ by a single

base pair substitution in this region and by a factor of two in the abundance of their mRNAs.

The egg shell protein genes, under study by A. Spradling (Carnegie Institution, Washington) and in F. C. Kafatos' laboratory (Harvard) also surprised us for they are amplified in the tissue in which they are expressed. Spradling has found that the amount of DNA amplified extends at least 30 kb on each side of the genes themselves, but the degree of amplification falls off, from a peak of some 60 fold, in both directions away from the coding sequences. This, and the fact that the amplified DNA has an identical restriction pattern to non-amplified DNA, may indicate that the process is chromosomal, rather than extra-chromosomal as in oocyte ribosomal genes in many organisms. Yet the Harvard laboratory (G. Thiroes and R. Griffin-Shea) have found that amplification of another pair of egg shell protein genes is accompanied by their

premature transcription; but that these early transcripts are not translated and, indeed, decay as development gets under way. Only at the appropriate developmental stage several hours later does transcription resume to produce mRNAs that are subsequently translated.

What is the meaning of this? It implies some relationship between amplification and early transcription and also indicates the existence of a translational control at the early developmental stage.

Like all good meetings, at the end there were far more questions than answers. As is right and proper we were treated to many very preliminary data whose place in the jigsaw puzzle is yet undiscovered. The great success of the Kolymbari meeting was in bringing together molecular and classical *Drosophila* biologists to discuss their data from their different perspectives and we can only hope that some of today's mysteries will be clearer to us when we next meet.

## Comets and the origin of life

from N.J. McNaughton and C.T. Pillinger

COMETS are the most spectacular bodies in our solar system and the anticipated return of Halley's Comet in 1985-6 has already proved a catalyst for increased cometary research. As the comet nucleus contains, perhaps with the exception of phosphorus, the necessary ingredients for life on this planet, the pundits of earthbound chemical evolution and cometary science found common ground at a recent colloquium\* entitled 'Comets and the origin of life'.

Although almost all of our observations of comets have been obtained from or near the Earth, the accumulated data yields a basic understanding of the structure, chemistry and physics of the comet nucleus and tails. Perhaps the most significant point to emerge from new observations is the apparent constancy of comet's chemical composition.

A'Hearn (University of Maryland) observed that the rate of production of certain chemical species, particularly CN, C<sub>2</sub>, C<sub>3</sub> and OH, from the comet nucleus near the sun is uniform irrespective of the type of comet and its gas to dust ratio. The spectroscopic continuums of comets of different ages are also strikingly similar (Donn, NASA-Goddard). It is not yet clear whether this uniformity is due to the nature of the chemical reactions in the coma or if it implies that comets are derived from a constant composition reservoir. How the latter interpretation fits with a recently announced theory for the origin of comets (Biermann *Royal Society Meeting on Planetary Exploration*, Munich, Nov. 4-5, 1980) is not clear. Biermann's new hypothesis predicts cometisimals as the

possible end product of a fragment of protosolar nebula collapsing in a relatively low magnetic flux. According to the same theory, planetary systems would be produced in a large magnetic flux and there seems to be no way for both planets and comets to be produced together.

The current interest in comets has promoted a shift in emphasis among the prebiotic synthesis investigators — low temperatures and photon or proton irradiation appear to be the order of the day. Reproducing cometary compositions and environments in the laboratory, the Maryland group and Greenberg (Leiden University, Netherlands) have synthesised 'life-like' organic molecules. Long chain hydrocarbons, nitrogenous compounds and carboxylic acids have been identified spectroscopically among the products of reactions carried out at temperatures as low as 15°K. Given that similar molecules could only be produced on the Earth under a strongly reducing protoatmosphere (which the Levine group from Virginia calculate would be short-lived due to the high reactivity and abundance of OH), the introduction of such molecules via comets into a less drastic protoatmosphere offers a real, and in some cases preferred, mechanism for the beginning of life on Earth. Speculations on this theme range from the controversial — Wickramasinghe (University of Cardiff) envisaging comets as the carriers of already living organisms — to efforts to calculate possible mass inputs (Lizcano-Aranjo, Houston).

Definitive evidence on the relationship between comets and the origin of life must

await the eventual laboratory analysis of cometary material. The identification of microscopic stratospheric and deep-sea particles of extraterrestrial origin by Brownlee (University of Washington) offers, however, the first real samples of probable cometary debris. The minute size and rarity of unaltered Brownlee particles preclude the use of most conventional analytical procedures, and substantial refinements to existing or even new techniques (such as the CD<sub>4</sub> carbon isotopic method pioneered at the University of Cambridge), are required before even the Brownlee particles can be fully exploited.

The alternative to analysing cometary material found in the terrestrial environment is *in situ* measurements from spacecraft. The planned NASA/ESA international mission to rendezvous with Halley's Comet which would have allowed analyses as sophisticated as those needed for origin of life studies will now not take place. ESA's fast flyby of Halley (the Giotto mission) will undoubtedly produce information on the composition of the gas and solid phases to fuel further laboratory based studies but is unlikely to provide definite answers to the question of whether comets were involved in the origin of life on Earth.

### CORRIGENDUM

Edward Herbert would like to point out that he was not the sole author of *Biosynthesis and processing of cellular and viral polypeptides* (*News and Views*, 13 November, 288, 115, 1980) and would like to acknowledge the co-authorship of Gunther Kreil (Salzburg) and the help of Carolyn Carter and Eckard Wimmer (Stoneybrook, New York).

N.J. McNaughton and C.T. Pillinger are in the Department of Earth Sciences, University of Cambridge.

\*The Fifth College Park Colloquium on Chemical Evolution was organized by Cyril Ponnamperna and held at the University of Maryland, Oct 29-31, 1980.