

*Nature conference*

# Thirty years of DNA

MOLECULAR biology seems, no doubt briefly, to be caught in the trap that gave early nineteenth century biology a bad name. The recognition that genes in higher organisms are interspersed with as yet apparently functionless stretches of DNA called introns seems to have stimulated a self-sustaining preoccupation with the most weird and the most wonderful. The piece most obviously missing from this jigsaw puzzle is a confident understanding of how genes became the kind of things they are, and how they control themselves (and each other).

This is one reaction to the conference to celebrate the 30th anniversary of the publication by J.D. Watson and F.H.C. Crick (*Nature* 171, 737; 1953) of a structure of DNA whose correctness has since been amply confirmed. The conference, organized by *Nature*, was held at Churchill College, Cambridge, on 13–14 April.

On such occasions, nevertheless, recollections of the past will out, advice to speakers notwithstanding. One confessed that he had been an intended mathematician when the Watson and Crick paper was published (W. Bodmer, Imperial Cancer



*Bodmer and Williamson*

Research Fund), another that he was anxious to be a chemical engineer (R. Williamson, St Mary's Hospital Medical School). Williamson was one of several European participants who drew attention to the time-lag in accepting the view that the structure proposed for DNA might be significant, quoting the then head of a British biochemistry department as saying that "this nucleic acid stuff will come to nothing".

Two session chairmen — Bodmer and Williamson — gave convincing reasons why molecular biology is, temporarily, natural history. Bodmer, once a student with R.A. Fisher, drew attention to the importance and antiquity of the concept of the gene cluster, defined by classical geneticists, but which is now quantifiable. In human genetics, the classical cluster is that of the globin genes, but now there are more telling systems such as those of the histocompatibility genes (HLA in human beings, H-2 in mice) and the immunoglobulin genes. If half of the  $3 \times 10^9$  base pairs of the human genome consists either of "junk" or of selfish DNA, and if an average cluster contains about 15 genes,

the problem of the human genome may be simply that of 1,000 to 3,000 gene clusters. The fact that actin-like genes have recently been found not always to specify proteins found in muscle seemed to Bodmer to be an illustration of how gene clusters would be defined by similarities of nucleotide sequences.

Multigene families cropped up repeatedly last week, most naturally in the summary by Lee Hood (Caltech) of what is known of the arrangement of genes in the 3 million base-pair region of mouse chromosome 17 which carries the genes for the type I and type II histocompatibility antigens as well as the genes encoding complement components, now known to be analogous to the HLA system in humans.

The rearrangement of the antibody genes is by now, of course, a familiar topic (see, for example, S. Tonegawa, *Nature* 302, 575–581; 1983). By supposing that the bits and pieces of the immunoglobulin genes on chromosomes 6, 12 and 16 in the mouse are rearranged in the lymphocytes that produce antibodies, and allowing for the occurrence of somatic mutation, it is possible to account for the capacity of the immune system to manufacture antibodies specific against an open-ended set of antigens: permutations available from among the several versions of each of several differently placed partial genes can easily explain 100,000 different antibody molecules even before allowing for mutation. The distinguishing feature of the family of antibody genes is their ability to undergo rearrangements which are mediated by conserved flanking sequences on either side of each of the pieces that eventually assemble to form a complete antibody gene. Could it be, asked Hood, that the similarities between these flanking sequences and those of "transposons" signal an evolutionary origin in migratory DNA, and the later exploitation of the migratory mechanism in the generation of antibody diversity? The source of the diversity of the transplantation antigens is, however, quite different, and consists of polymorphism.

Hood's purpose, last week, was to explain how this genetic diversity might have come about. In the mouse major histocompatibility complex (MHC) many more gene-like sequences have been detected than are needed to account for the known

*Hood and Borst*



*Jim Watson, in his introduction to the conference, looked back . . .*

WHY thirty years and not, say, twenty years or ten years? Well, it's pretty clear why it couldn't be ten years, because we waited  $9\frac{1}{2}$  years to get the Nobel Prize. Fifteen years wouldn't have had any meaning, as the big event occurred after thirteen years, in 1966, when the genetic code was solved. Then repressors came along, but that wasn't enough for a big celebration. And then twenty — we didn't know it at the time, but that was the real year, because that was when we came to recombinant DNA. It was made practical by Boyer and Cohen. People ask me "Why did the Swedes take nine years to give you the Nobel Prize if it was so important?" But they still haven't given it to Boyer and Cohen.

*Nature* did celebrate twenty-one years, and that was really very nice, because Francis [Crick] wrote and, in particular, Linus [Pauling] wrote. But we couldn't sell the twenty-fifth anniversary very big, because we were still mad at each other. We couldn't work on tumour viruses, because of the regulation of recombinant-DNA research. And it's really only now, in year thirty, that we're free to do exactly as we want.

Now, I guess, people like me have lots of debts to acknowledge — mostly to people that those here really didn't know. But there were two unique patrons — people who protect you if you're trying to do something different. One was the Rockefeller Foundation. It really started the thing. It gave money in 1933 to begin the first Cold Spring Harbor Symposium, and it gave a lot of money to Caltech and it gave money to the Medical Research Council (MRC). It gave money to Indiana University. It gave money to Copenhagen, when I was with Herman Kalckar. The Rockefeller Foundation did this because of one man, Warren Weaver, a mathematician from the University of Chicago, president of the University of Wisconsin, and . . . a

fine man. It was his belief that helped Linus Pauling and George Beadle to talk about how chemistry was going to revolutionize biology. Weaver provided a spirit in which younger people could think.

The second patron was MRC. Harold Himsworth supported the MRC laboratory in its earliest days and Sidney Cattell, my better, saw that the new laboratory was built, without which Cambridge couldn't have been the outstanding place it still is today.

The second kind of patron is more personal. We had Sir Lawrence Bragg. As the Cavendish Professor, he could assign space. And he assigned some space to Max Perutz and then to John Kendrew so that there were six people in Cambridge when I arrived in 1951. And, as I recount in *The Double Helix*, I thought Bragg was just a stuffy old man when I met him. But he was a fine man. He had a really keen interest in science, and he was certainly Francis's only competition at that time — in the sense that he was a theoretician. And he had a difficult time, because most people thought that it was his father who had been the clever one, whereas it was the younger Bragg who'd made the running. When he came to the Cavendish, people said how dreadful it was that he was not a nuclear physicist or even a low-temperature physicist, but just a crystallographer. So his support of us was very important. I showed him several manuscripts, and he really was very helpful.

Then there was the patronage given to Francis and me by John Kendrew and Max Perutz. Francis was Max's research student and I was John Kendrew's. And when I was going to lose my fellowship for coming to Cambridge, they dug up enough money for me to stay in Cambridge.

But I guess I owe most of all to Francis, who really did look after me, and who often tried to keep me from being silly. I wasn't as silly as he thought, but he was so sensible that I had occasionally to say things I didn't believe, to see if I could trap him. And I sometimes did.

I don't think the whole thing would have worked if the Cricks hadn't cooked so many meals for me, or made me feel at home in Cambridge by seeing that I didn't cut my hair for quite a while. And then seeing that I wore a tie. And that I got an English suit. And giving me the good advice that I shouldn't look like an American. I followed their rules to the point where it made it difficult sometimes to go home.

I couldn't have got anywhere without Francis, so I really felt a little strange coming back here for this meeting, because it's without Francis. It could have been Crick without Watson, but certainly not Watson without Crick.

I wrote *The Double Helix* because it was a good story. Some people claimed they didn't recognize me in it, but others thought they recognized me too well. Originally, it had the title *Honest Jim*, but some people objected because they thought

polymorphic class I antigens and their non-polymorphic relatives, the lymphocyte differentiation antigens. This has given rise to the suggestion that the extra gene-like sequences in the MHC may represent a reservoir of genetic information that could contribute to the polymorphism.

Hood also advised looking for multigene families elsewhere. Will some gene cluster, or a few, explain the specificity of, for example, the detection of smell. Maybe. But can it be so simple?

This interest in movable genes, the malleable genome, is now more widely shared. The most striking illustration, last week, of the importance of the malleability of the genome was the account by Pro-



Hogness and Temin

fessor P. Borst (Amsterdam) of the mechanism of antigenic variation by trypanosomes, the bloodstream parasites responsible for sleeping sickness. In the course of a trypanosome infection, as many as a hundred changes may be rung on the antigen specificity of the protein component of the outer envelope. Hybridization studies involving cDNA and native DNA are consistent with the view that trypanosomes carry a stock of silent genes which spring into action only when they migrate to the neighbourhood of what Borst calls "an expression site".

How is this accomplished? As Borst explained, it appears that there is more than one way of getting a new antigen gene into the "expression site", which is right at one end of a trypanosome chromosome. One possible mechanism is reciprocal exchange with the end of another chromosome, replacing the expressed antigen gene but leaving the rest of the chromosome intact. To have the requisite number of silent genes at these ends of chromosomes, trypanosomes carry several hundred "mini-chromosomes".

As it happens, trypanosomes have only six times as much DNA as *Escherichia coli*, much of which must be given over to the genetic ingredients of antigenic variability (but trypanosomes cannot live by agar alone). *Drosophila melanogaster* is, by contrast, more like the human genome in its complexity, with the further attraction



Jaenisch and Klug

of carrying known genes which regulate the development of the organism by some means still unknown. David Hogness described the success with which his colleagues at Stanford have nevertheless been able to clone the huge 75,000 base-pair region of the *Drosophila* genome in which homeotic mutations of the bithorax complex of the *Ultrabithorax* class are known to be scattered. The surprise is that most of the spontaneously arising mutations appear to entail the insertion of transposable elements into two short segments towards the ends of the 75,000 base-pair transcriptional unit that apparently encode the major exons of the discrete RNA products of this gene. The sheer size of this transcription unit remains something of a puzzle, and Hogness wondered whether it could be a mechanism for timing developmental events.

Movable genetic elements were also conspicuous parts of last week's account by Howard Temin (University of Wisconsin) of the molecular biology of the retroviruses capable of transforming cells into a malignant condition. That RNA can carry genetic information is well established. Temin asked his audience to consider exogenous retroviruses as being in a kind of equilibrium with the corresponding genetic information (written in DNA) in the genome of the infected cell, which is in turn potentially a movable element.

Temin's own work with the only retrovirus known to infect turkeys has persuaded him of the complexity of sequence of genetic transformations that must precede the conversion of cellular genetic information into the viral oncogene, involving the fashioning of a new promoter and the beginning of the gene, the elimination of intervening sequences and a variety of point mutations. Perhaps because of these changes, when the viral oncogene is integrated into the cellular genome, it is at a chromosomal site different from that of the progenitor gene. The problem of the genetics of cancer, he said, is far from understood.

R. Jaenisch (Hamburg) has used retroviruses for a different purpose, to discover into which embryonic genes they are inserted in such a way as to create a lethal, recessive mutation. By breeding experiments in which hundreds of mice have been involved, Jaenisch has so far discovered one such insertion which, in homozygous mice, produces intrauterine death at about day fifteen.

Why should that be? Jaenisch appears to have had the good fortune that the viral gene in the affected animals had been inserted in a region of the mouse genome coding for one type of collagen. As it happens, the viral gene is incorporated backwards, at the beginning of the collagen gene, which may be one reason why homozygous embryos fail to produce the protein. But Jaenisch says that this cannot by itself be a complete explanation.

In the natural history of genes, catalogu-

ing occurrences along linear stretches of DNA is the obvious first step. Telling how the DNA is disposed in chromosomes has turned out to be a harder task, but the conference last week was told (by Aaron Klug, MRC Laboratory of Molecular Biology, Cambridge) of new X-ray diffraction data from single crystals of nucleosomes — units of chromatin consisting of 140 base pairs of DNA wound in  $1\frac{3}{4}$  turns of a left-handed superhelix around a core of eight protein histone molecules.

The Cambridge group has obtained X-ray diffraction data at a resolution of 7 Å, at which point the disposition of the two grooves (major and minor) of the helix can be identified. The windings of the molecule on the histone core are closer together than expected. The problems of understanding how such compact structures are replicated and transcribed remain.

Another way of understanding how DNA molecules dispose themselves in real life, described last week by J. Wang (Harvard), smacks more of physical chemistry. Take a circular molecule of double-stranded DNA from polyoma virus, and artificially insert bits and pieces of DNA so as to tell (from the mobility of the molecules thus formed) how the degree of supercoiling has been affected.

Wang described how stretches of the repeated C-G (cytosine, guanine) sequence do indeed affect the mobility of the circular molecules in a way that suggests that a piece of Z-DNA (left-handed) has been formed. He has also been able to calculate the free energy of formation (60 kcal per mole) of cruciform structures formed by the insertion of an artificial sequence of 68 base-pairs. Wang's plan now is to study the physiological significance of supercoiling with an artificial plasmid containing the gene for topoisomerase from *E. coli*.

Mere description, unfortunately, does not explain how genes may actually function. The best known piece of naturally occurring genetic machinery remains the working of the lambda repressor, the gene carried by the lambda-phage which, in *E. coli*, can turn off the genes responsible for replication of the virus. The principal present custodian of this piece of machinery, M. Ptashne (Harvard) described how the N-terminal ends of pairs of repressor molecules attach themselves to one of the three contiguous sites in the major groove of a DNA helix so that part of the attaching protein embraces the whole helix.

By model-building and with the help of lambda mutants whose structural consequences can be predicted, a virtually complete picture of the stereochemistry of repressions has been built up. Ptashne says that his immediate ambitions include making mutants of the bacterial RNA polymerase to confirm that repressor molecules control its activity in the way now predicted, to tinker with the binding sites of the repressor molecules and to study newly obtained co-crystals of the repressor of bacteriophage 434 and

bacterial DNA.

The functions (if such there are) of intervening sequences remain unclear. How, in any case, did they arise? C. Weissmann (Zurich) gave it as his opinion that there is no way of telling whether genes with introns are more or less "primitive" than those without. Virus and bacteria lack them, but they survive.

The stability of introns within particular genes argues for function (although the frequent appearance of particular sequences within introns is more probably necessary for the accuracy of splicing). Weissmann took the view that introns are very old, perhaps parts of the earliest RNA genomes, and that they must during the



Ptashne and Wang

evolution of different genomes have been a valuable means of amplifying functional genes and of making possible genetic rearrangements without breaking up coding sequences of DNA. How else could the family of immunoglobulin genes have arisen? But with all this said, the complexity of the collagen gene remains a puzzle. Perhaps (as Weissmann said in another context) "Nature is trying to tell us something".

W. Gilbert (Biogen) put forward last week another way of coming to grips with the same question — the notion that the functional parts of split-up genes have indeed been recruited from elsewhere, and that as a consequence the distribution of introns should be reflected in the structure of the proteins eventually produced.

The notion that protein molecules consist of recognizably distinct domains is not of course new. Gilbert, however, was able last week to show that the view has some predictive value, in spite of some disconcerting evidence — the recognition, for example, that the eight-fold symmetry of the enzyme triose phosphate isomerase (from the chicken) is coded for by seven and not eight gene pieces.

To describe all this as natural history is, of course, unfair. For the phrase applies only in the sense that while genetic maps are still incomplete, most new information starts fresh hares. And there have already been several telling illustrations of how it is

Weissmann and Gilbert



I was proclaiming myself as the Honest Truth. So then I changed it to *Base Pairs*, and that didn't go over either, so the book ended up with the title *The Double Helix*. I had this great idea of a picture of me looking at Francis and Maurice [Wilkins] looking at Rosalind [Franklin] on the cover of *Base Pairs* — a kind of *Kind Hearts and Coronets* arrangement. But it ended up a little more sensible.

Rosalind Franklin was a very intelligent woman, but she really had no reason for believing that DNA was particularly important. She was trained in physical chemistry. I don't think she'd ever spent any length of time with people who thought DNA was important. And she certainly didn't talk to Maurice [Wilkins] or to John Randall, then the professor at Kings. And then the time came when he moved her out of DNA and sent her over to Bernal, and she had to write up the papers.

As Aaron Klug has explained in *Nature*, she came very close to the structure of DNA. She really had accepted that it was a helix, but we didn't know that. She was moving towards the idea that there were two chains, but she never built models. Now if she could have had just two hours with Francis, and she could have been convinced that he was right, not just a loud-mouth, I think she would have gone back to her laboratory and built models. She'd have solved the structure of DNA. But she was really prepared to give up working on DNA, and she wouldn't have agreed to give up if she'd thought it was important. So that was why she didn't get the answer.

Probably none of you knows how much she liked Francis. Afterwards she came to him to talk. She did that very beautiful work on tobacco mosaic virus (TMV) and when she had her operation for ovarian cancer, which she knew was very serious, she came and stayed with the Cricks to recuperate. She was pretty ill, but she was supported strongly by Don Caspar, and less so by me, and we got her a National Institutes of Health grant that let her do the TMV work when the Agricultural Research Council (ARC) turned her down after she'd told the head of ARC that he was an idiot. He was an idiot. When she wanted a diffractometer, he pointed out that there was one in Aberdeen. So she told him what an idiot he was and then they didn't support her. She wasn't a diplomatic person but when you got to know her she was fine to deal with.

Linus [Pauling] didn't deserve to get the structure. He really didn't read the literature. And he didn't talk to anyone either. He'd even forgotten his own paper with Max Delbrück which said that a gene should replicate by complementarity. He seems to consider that he should have got the structure because he was so bright, but really he didn't deserve it.

Now he might have got it because Alex Rich arrived at Caltech and began to take X-ray photographs at just about the time that we proposed our model. I think it was

inevitable that the structure would have been solved within about a year. The momentum was there, and they knew that DNA was important.

But I have no guilt feelings. Both Francis and I were products of a tradition that wanted to solve the problem, right through. I'd been trained in the phage group, for which self-replication was the key problem. And Francis was the heir to the tradition established here in England that molecular structure is important, and if you work out bigger and bigger structures, you'll learn something important.

What happened when we got the structure? Well, we wrote the paper for *Nature* and then we realized that the final throw-away line might have been too cute and that we'd better write a longer paper. So, so that some third person wouldn't write down these trivial ideas and claim credit for them, we wrote the second *Nature* paper called "The genetical applications of the structure of DNA". And then Francis started to talk. I mean he'd always talked. And then Sydney [Brenner] came over. On about the first occasion I saw Sydney, we talked for about six hours non-stop. Sydney had a few bright friends in Oxford who talked about DNA, but they didn't have expert pictures and didn't do anything about it. Sydney was then a research student with Hinshelwood, who was a kind of Lamarckian and . . . a strange man. And Sydney was doing a PhD thesis that was as dull as Francis's, on bacteriophage T4. But he knew the phage was important.

He took me aside one day and said "Jim, you don't realize how important the work you've done is". I think I did, but I was also scared — it might not be right. One had this sort of feeling because of the Cambridge biochemists who were calling it the WC structure. When Francis got a request to speak on BBC radio, I said that he could do it if the broadcast went out on the External Services and if it wasn't heard in England. So his voice talking about DNA and its importance was heard only outside England. But then I realized that it was the end of an era. I'd been raised at courses on H.J. Muller, and I'd heard what genes might be and how they might self-replicate. So when we saw the complementary structure of DNA, that was the final solution. We didn't see it, as we now see it, as the beginning.

Both Francis and I had no doubts that DNA was the gene. But most people did. And again, you might say, "Why didn't Avery get the Nobel Prize?" Because most people didn't take him seriously. Because you could always argue that his observations were limited to bacteria, or that [the transformation of *Pneumococcus* that he described was caused by] a protein resistant to proteases and that the DNA was just scaffolding. Only when they saw the double helix did people stop asking what the gene was.

The first thing I worried about was whether the strands would come apart

quickly enough and how to explain [Arthur] Kornberg's discovery that the chains run in opposite directions. But I doubt whether Francis and I combined spent more than ten hours worrying whether the structure was right, whether they would take it away from us or something like that. So it was very satisfying when we saw that the replication scheme seemed to be right, even though [L.F.] Cavalieri came along and said we should use a four-chain model.

What happened next? There was the area of the genetic code, or how information was included in DNA. That was dominated by Francis and Sydney, Khorana and a few others, but I'd say the work here at Cambridge was the intellectual high-point of that whole thing. Then we were all dominated by the biochemistry of the central dogma [DNA makes RNA makes protein] and we all wanted to find the enzymes that did all those steps. That was an era when we were the new enzymologists. But when Francis and I saw the structure, we'd never thought in terms of an enzyme. We even thought that because it's a template, you might not need those enzymes. But now, of course, there's this great RNA splicing story without enzymes that makes one think that we weren't totally mad. But it was far out, and the person who used to argue with us most about enzymes was Peter Mitchell, which was interesting because [in his later work] Peter managed to do away with the need for several enzymes.

Without first class biochemistry of this sort, from first class conventional biochemists such as Lippman and Ochoa, we wouldn't be here today, because the enzymes were necessary. But then we also needed the wonderful molecular genetics which came from people such as Benzer, Sydney [Brenner], Jacob and Monod and which finally led to Mark [Ptashne] and Wally [Gilbert] and the repressor story.

But then we didn't know what to do next. Many of us moved into tumour viruses, hoping that what had worked for phage and bacteria would work for tumour viruses in animal cells. It's chilling to think of where we'd be if the cloning of DNA hadn't come along. We would be stuck so far back that it wouldn't be much fun. Even there we owe an enormous amount to a genetics approach that has really made us think about plasmids and so on until, finally, recombinant DNA came along.

But then we were involved with the question of whether we had something so powerful that we shouldn't do it. We'd gone through that. And [the fuss about the control of recombinant-DNA research] is now so far back that I don't really feel mad. I always regarded the regulation of recombinant-DNA research as a black comedy, and said so. But some people thought the concern was genuine. It was genuine for some, but for others it was really the theatre of the absurd. And it's over. What a relief! □



Nasmyth and Jackson

possible deliberately to explore the genome for particular functions.

Much in this spirit, J. Jackson (Columbia) is one of the growing band embarking on the exploration of the nervous system. He described work in which he and his associates have identified and cloned genes encoding three previously identified neuropeptides that control the behaviour at egg-laying of the marine invertebrate *Aplysia*. These genes generate polypeptides that can be enzymatically processed to yield not only the three isolated peptides but, potentially, a host of others.

Gene regulation is another preoccupation. Last week, K. Nasmyth (MRC, Cambridge) described his hunt for the means by which the switch from one mating type to another is accomplished in yeast. H. Pelham from the same laboratory demonstrated the power of the techniques now available for incorporating genes into animal cells (HeLa cells and *Xenopus* oocytes, for example) as a means of identifying control elements upstream of the point at which transcription begins.

Biotechnology apart, Y.W. Kan (San Francisco) gave an account last week of how molecular biology can prevent genetic disease, by the use of genetic probes to detect patterns of inheritance.

So where will it lead? Sydney Brenner (MRC, Cambridge) offered a tantalizing vision of what may lie ahead. There may be Pelham, Brenner and Kan



no commanding heights still to be scaled, but only an apparently endless succession of foothills. Perhaps, "when we've done thymidine kinase and all the other enzymes we can think of", there will be little left.

So is there a magic room in which the secrets of life will be found? "We open one door, and find ourselves in a long corridor, at the end of which there is another door, and then another corridor, and so on." That, Brenner said, is what science is like. The next breakthrough? The question may be too late. Perhaps the breakthroughs have already been made and we need not await another Galileo, "the Watson and Crick of development", Brenner said. "Maybe that's what it's really like. People will just publish more and more . . ." □