Towards a paradigm shift in biology

The steady conversion of new techniques into purchasable kits and the accumulation of nucleotide sequence data in the electronic data banks leads one practitioner to cry, "Molecular biology is dead — Long live molecular biology!".

THERE is a malaise in biology. The growing excitement about the genome project is marred by a worry that something is wrong — a tension in the minds of many biologists reflected in the frequent declaration that sequencing is boring. And yet everyone is sequencing. What can be happening? Our paradigm is changing.

Molecular biology, from which has sprung the attitude that the best approach is to identify a relevant region of DNA, a gene, and then to clone and sequence it before proceeding, is now the underpinning of all biological science. Biology has been transformed by the ability to make genes and then the gene products to order. Developmental biology now looks first for a gene to specify a form in the embryo. Cellular biology looks to the gene to specify a structural element. And medicine looks to genes to yield the body's proteins or to trace causes for illnesses. Evolutionary questions — from the origin of life to the speciation of birds — are all traced by patterns on DNA molecules. Ecology characterizes natural populations by amplifying their DNA. The social habits of lions, the wanderings of turtles and the migrations of human populations leave patterns on their DNA. Legal issues of life or death can turn on DNA fingerprints.

And now the genome project contemplates working out the complete DNA pattern and listing every one of the genes that characterize all of the model species that biologists study — ourselves even included.

At the same time, all of these experimental processes — cloning, amplifying and sequencing DNA — have become cook-book techniques. One looks up a recipe in the Maniatis book, or sometimes simply buys a kit and follows the instructions in the inserted instructional leaflet. Scientists write letters bemoaning the fact that students no longer understand how their experiments really work. What has been the point of their education?

The questions of science always lie in what is not yet known. Although our techniques determine what questions we can study, they are not themselves the goal. The march of science devises ever newer and more powerful techniques. Widely used techniques begin as breakthroughs in a single laboratory, move to being used by many researchers, then by technicians, then to being taught in undergraduate courses and then to being supplied as purchased services — or, in their turn, superseded.

Fifteen years ago, nobody could work out DNA sequences, today every molecular scientists does so and, five years from now, it will all be purchased from an outside supplier. Just this happened with restriction enzymes. In 1970, each of my graduate students had to make restriction enzymes in order to work with DNA molecules; by 1976 the enzymes were all purchased and today no graduate student knows how to make them. Once one had to synthesize triphosphates to do experiments; still earlier, of course, one blew one's own glassware.

Yet in the current paradigm, the attack on the problems of biology is viewed as being solely experimental. The 'correct' approach is to identify a gene by some direct experimental procedure — determined by some property of its product or otherwise related to its phenotype — to clone it, to sequence it, to make its product and to continue to work experimentally so as to seek an understanding of its function.

The new paradigm, now emerging, is that all the 'genes' will be known (in the sense of being resident in databases available electronically), and that the starting point of a biological investigation will be theoretical. An individual scientist will begin with a theoretical conjecture, only then turning to experiment to follow or test that hypothesis. The actual biology will continue to be done as "small science" — depending on individual insight and inspiration to produce new knowledge — but the reagents that the scientist uses will include a knowledge of the primary sequence of the organism, together with a list of all previous deductions from that sequence.

How quickly will this happen? It is happening today: the databases now contain enough information to affect the interpretations of almost every sequence. If a new sequence has no match in the databases as they are, a week later a still newer sequence will match it. For 15 years, the DNA databases have grown by 60 per cent a year, a factor of ten every five years. The human genome project will continue and accelerate this rate of increase. Thus I expect that sequence data for all of the model organisms and half of the total knowledge of the human organism will be available in five to seven years, and all of it by the end of the decade.

To use this flood of knowledge, which will pour across the computer networks of the world, biologists not only must become computer-literate, but also change their approach to the problem of understanding life.

The next tenfold increase in the amount of information in the databases will divide the world into haves and have-nots, unless each of us connects to that information and learns how to sift through it for the parts we need. This is not more difficult than knowing how to access the scientific literature as it is at present, for even that skill involves more than a traditional reading of the printed page, but today involves a search by computer.

We must hook our individual computers into the worldwide network that gives us access to daily changes in the database and also makes immediate our communications with each other. The programs that display and analyse the material for us must be improved — and we must learn how to use them more effectively. Like the purchased kits, they will make our life easier, but also like the kits, we must understand enough of how they work to use them effectively.

The view that the genome project is breaking the rice bowl of the individual biologist confuses the pattern of experiments done today with the essential questions of the science. Many of those who complain about the genome project are really manifesting fears of technological unemployment. Their hard-won PhDs seem suddenly to be valueless because they think of themselves as being trained to a single marketable skill, for a particular way of doing experiments. But this is not the meaning of their education. Their doctorates should be testimonials that they had solved a novel problem, and in so doing had learned the general ability to find whatever new or old techniques were needed; a skill that transcends any particular problem.

Walter Gilbert

Walter Gilbert is Carl M. Loeb university professor in the Department of Cellular and Developmental Biology at Harvard University, in the Biological Laboratories, 16 Divinity Avenue, Cambridge, Massachusetts 02138 USA.