Genetics in the post-sequence era

The last decade of the twentieth century will see the dawn of a new age of genetics that I like to call the 'post-sequence era'. Towards the end of the decade, complete genomic sequences will become available for model organisms with small genomes: bacteria and yeast. Shortly thereafter we should have the sequences of the worm and fly genomes, and some time early in the next century we can expect the complete sequence of the human genome. The availability of complete genomic sequences will change, dramatically and permanently, the way in which we geneticists think and work. By relieving us of the need to focus on the mostly mechanical preliminaries to serious study of genes (mapping, cloning, and especially sequencing), access to complete genomic sequences will allow us to concentrate on their biological functions.

It is a human trait, especially of those who are strongly goal-oriented, to make light of, even to disregard, the pain and difficulties prerequisite to achievement. Like athletes, productive scientists understand that world-class achievements do not come without world-class effort. I think this is why molecular geneticists have generally played down and understated the cost and effort that now goes into cloning and sequencing. My informal estimate is that as much as one third of all the money and effort now being spent in molecular genetics goes on what one might call cloning and sequencing 'overheads'. Worse, this preliminary cloning and sequencing diverts attention from the real scientific issues that can only be addressed after we know our sequences.

Complete genomic sequences will free us from both the overheads and the diversion. By the year 2000, geneticists working with bacteria (*Escherichia coli* and maybe others like *Mycobacterium tuberculosis*) and yeast (*Saccharomyces cerevisiae* and perhaps *Schizosaccharomyces pombe*) will no longer have to suffer the manual, *de novo* sequencing of one or more genes for each and every project. This rite of passage will join other bygone rites (like purifying one's own *Bam*HI or blowing one's own glassware) in the dustbin of technological history. Progress in understanding biological functions will accelerate as biologists in the post-sequence era study and manipulate more than just one or two genes at a time.

A specific example might help to illustrate my point. Suppose that in the future a student devises a screen for a novel phenotype as a way of finding genes involved in a particular function. The student will be able to draw not only on the complete sequence, but also on the materials generated as prerequisites to the organized sequencing of any genome, such as arrayed clone banks with complete coverage of the genome. This will allow mapping of new mutations by complementation. Simple experiments will localize the mutations to small (less than a few hundred base pair) segments that can be recovered using the polymerase chain reaction (PCR). The gene in which the mutations lie will then be known for certain; a single sequence run will suffice to characterize such a mutation completely. With very few experiments, even an absolute novice will be able to get to the biological issue of how the gene acts to produce the phenotype.

In the future, genes will be identified not only through mutation or via the protein but, more often, directly from the sequence. It will be possible to recognize all members of a gene family by sequence comparison in the time it takes a computer to search the databases. This method is much quicker, and indeed more reliable, than hybridization and PCR techniques. Knowing the sequence of every gene, it will be straightforward to construct perfect null mutations. When such a null mutation has no phenotype, it will be easy to combine it with null mutations in closely homologous genes.

Research on the model organisms whose genome sequence is known will be markedly accelerated. Attention will shift from the identification of the genes and gene products involved in a process like mitosis, to the mechanisms by which these genes and proteins function, and how they interact.

In conclusion, I expect that by the year 2000 all biologists will have come to recognize that complete genomic sequences are a vital tool for biology. Indeed, genomic sequences will have an influence comparable to instruments, like the microscope, the spectroscope and the X-ray diffractometer, that in their day revolutionized science by allowing us to see our subject in a completely new way.

DAVID BOTSTEIN

Department of Genetics, Stanford University School of Medicine, Stanford, CA 94305, USA.

Any Technical Tips?

Technical Tips is a place where readers can exchange information about useful lab techniques. Technical Tips should be either new methods, or significant new modifications or applications of existing techniques. If you have developed a handy new method, why not share it with other *TIG* readers? Your article should be as brief as possible, but should give enough information to enable others to repeat the method. If any part of the method involves published procedures, you can refer to the appropriate paper(s) rather than repeating those details. All Technical Tips are peer-reviewed.

Please send three copies of your double-spaced typescript, plus three copies of any figures (including at least one set of originals) to:

Dr Alison Stewart, *Trends in Genetics*, Elsevier Trends Journals,
68 Hills Road, Cambridge, UK CB2 1LA.