PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Preface

Walter Bodmer and Peter Donnelly

Phil. Trans. R. Soc. Lond. B 1994 **344**, 327
doi: 10.1098/rstb.1994.0070

PREFACE

The technological revolution in molecular biology over the past 10–15 years has opened vast new horizons for exploration. It has also dramatically increased the amount of information available on organisms at the molecular level. The interpretation of this new information, and its management and the design of the experiments which lead to it, has in turn raised challenging problems. Often, mathematical and statistical ideas have been indispensible to progress. As the papers in this volume show, the interaction is not confined to one particular area of the mathematical sciences. In some settings, existing results have been ideally suited to the biological problem. In others, progress has itself stimulated important mathematical advances.

The human genome project has provided substantial impetus to this endeavour. Physical mapping involves the construction of an ordered library of 'clones', overlapping pieces of DNA of convenient size, which covers the genome. Several different experimental strategies are in use, each of which gives rise to a range of mathematical, statistical and computational problems. The aim of genetic mapping is to specify the relative locations of genes and genetic markers on chromosomes. New methods for defining polymorphic markers have coincided with major advances in computational statistics, so that the analysis of data from pedigrees of reasonable size, for realistic genetic systems, is now possible.

Many of the mathematical and computational problems in the area have their origins in DNA sequence data itself. Examples include the study of sequence structure and composition, the optimal alignment of sequences from different genes and organisms, and the detection of coding regions and other interesting features in a sequence.

One particularly important problem that has attracted considerable attention is the use of the DNA (or amino acid) sequence in a coding region to predict the three-dimensional structure of the associated protein. Some progress has been made here (and elsewhere in the molecular context) through the use of techniques from artificial intelligence. It remains the case, however, that biochemical insight and further experimental work will form an essential part of future developments.

In addition to the direct information relating DNA sequences to their functional roles, the growth in molecular data provides valuable information with which to study evolution. There are two different levels, and two different timescales, on which data is of interest. Information on sequence variability within a species can be used to infer aspects of evolution within that species. In some cases the evolutionary mechanisms themselves are of interest (estimates of mutation rates and selection intensities for example), whereas in others, the genetic data may carry information about the history of the species itself (for example, questions related to the size and distribution of ancestral populations are currently of substantial interest in connection with the human population). On the other hand, comparisons of sequence data between species are informative about evolution on a longer timescale, either in reconstructing the evolutionary relationships between the species involved or in studying the relevant evolutionary mechanisms.

One of our aims in organizing the discussion meeting was to bring together the various academic communities and to draw attention to the challenging and important problems which arise. The papers in this volume describe particular aspects of these problems. In addition, each serves to highlight the benefits of cross-disciplinary communication and interaction in a rapidly developing field.

We are especially grateful to the speakers at the discussion meeting for all their efforts, then and subsequently, and to all of those who participated in the meeting. We would also like to record our thanks to the Royal Society for its support for, and organization of, the meeting, and particularly to Mary Manning, Christine Johnson, Janet Clifford and Simon Gribbin for their work leading up to the meeting and in connection with the preparation of this volume.

May 1994

Sir Walter Bodmer
Peter Donnelly