

The applications of computers to research on nucleic acids II Part 2

**Edited by
D Söll & R J Roberts**

The applications of computers to research on nucleic acids II Part 2

**Edited by
D Söll & R J Roberts**

Originally published as
Nucleic Acids Research
Volume 12 Number 1, January 1984
and supplied to subscribers as part
of their 1984 subscription

ISBN 0-904147-59-2

 **IRL PRESS**
Oxford · Washington DC

IRL Press Limited,
O.Box 1, Eynsham,
xford OX8 1JJ,
ngland

© 1984 IRL Press Limited

All rights reserved by the publisher. No part of this book may be reproduced or transmitted in any form by any means, electronic or mechanical, including photocopying, recording or any information storage and retrieval system, without permission in writing from the publisher.

British Library Cataloguing in Publication Data

The applications of computers to research on
nucleic acids. II.

I. Nucleic acids—Research—Data processing

I. Söll, D. II. Roberts, R.J.

547.7'9'072 QD433

ISBN 0-904147-59-2

Printed in England by Information Printing Limited, Eynsham.

85 - 307
B1

PREFACE

Two years have passed since the first special issue of NUCLEIC ACIDS RESEARCH devoted to the applications of computers to research on nucleic acids (vol 10, 1-456, 1981). In that time many laboratories have acquired on-line access to their university computer or have invested in a personal microcomputer. No longer are these machines viewed as toys, but rather have assumed the importance of other, more familiar items of equipment in a molecular biology laboratory. Two major data bases, GENBANK and the EMBL data base, have come into existence in response to the need for computer access to published sequences. The software needed to analyze nucleic acid and protein sequences continues to be developed, and already one commercial enterprise has emerged that both develops and maintains a large software package for such analysis.

Despite the widespread interest in programs for manipulating nucleic acid sequences rather few of them are well known since most have been developed locally and do not enjoy widespread circulation. It is the aim of this second special issue to heighten the awareness of both scientists and programmers to the broad range of software that is currently available. It is our hope that this issue, like the previous one, will provide a useful starting point for those investigators newly discovering a need for programs in this area and will stimulate those already deeply involved to exchange and improve existing programs.

Richard J. Roberts

Dieter Söll

Contents

MINICOMPUTERS (continued)

A computer program for the management of small cosmid banks D.M.Lonsdale, T.P.Hodge and P.J.Stoehr	429
Computer programs in nucleic acid synthesis: synthetic strategy development using solid-phase chemical techniques with data storage, retrieval and analysis capabilities S.Lombardi, H.Seidell, S.Pulford, W.Dutton and S.Parekh	437
A fast homology program for aligning biological sequences P.Taylor	447
Sequence comparison by exponentially-damped alignment D.R.Boswell and A.D.McLachlan	457
PLASMAP: an interactive computational tool for storage, retrieval and device-independent graphic display of conventional restriction maps B.N.Stone, G.L.Griesinger and J.L.Modelevsky	465
Computer-based image analysis of one-dimensional electrophoretic gels used for the separation of DNA restriction fragments A.J.Gray, D.E.Beecher and M.V.Olson	473
GELYSIS: Pascal-implemented analysis of one-dimensional electrophoresis gels P.J.Russell, R.E.Crandall and R.Feinbaum	493
A computer program to enter DNA gel reading data into a computer R.Staden	499
Computer methods to locate signals in nucleic acid sequences R.Staden	505
Graphic methods to determine the function of nucleic acid sequences R.Staden	521
The codon preference plot: graphic analysis of protein coding sequences and prediction of gene expression M.Gribskov, J.Devereux and R.R.Burgess	539
Measurements of the effects that coding for a protein has on a DNA sequence and their use for finding genes R.Staden	551

MICROCOMPUTERS

Apple II Pascal programs for molecular biologists B.Malthiery, B.Bellon, D.Giorgi and B.Jacq	569
A comprehensive sequence analysis program for the IBM personal computer C.Queen and L.J.Korn	581
DNA sequence analysis on the IBM-PC W.F.Schwindinger and J.R.Warner	601
A DNA sequence analysis package for the IBM personal computer L.M.Lagrimini, S.T.Brentano and J.E.Donelson	605

Contents continued overleaf

A computer package for DNA sequence analysis F.R.Blattner and J.L.Schroeder	615
Update for users of the Cornell sequence analysis package J.S.De Banzie, E.W.Steeg and J.T.Lis	619
DNA sequence handling programs in Basic for home computers P.A.Biro	627
BIOLOG — a DNA sequence analysis system in Prolog A.Lyall, P.Hammond, D.Brough and D.Glover	633
A convenient and adaptable package of computer programs for DNA and protein sequence management, analysis and homology determination J.Pustell and F.C.Kafatos	643
JINN, an integrated software package for molecular geneticists M.Johnsen	657
Interactive computer programs for the graphic analysis of nucleotide sequence data V.A.Luckow, R.K.Littlewood and R.H.Rownd	665
An inexpensive semi-automated sequence reader for the Apple II computer M.Komaromy and H.Govan	675
DNAGEL: a computer program for determining DNA fragment sizes using a small computer equipped with a graphics tablet T.Kieser	679
Apple software for analysis of the size of restriction fragments J.I.Rood and J.M.Gawthorne	689
Molecular weight determination program C.V.Maina, G.P.Nolan and A.A.Szalay	695
An efficient program to construct restriction maps from experimental data with realistic error levels R.Durand and F.Bregegere	703
Plasmid mapping computer program G.P.Nolan, C.V.Maina and A.A.Szalay	717
KLONER: a computer program to simulate recombinant DNA strategies by restriction map manipulation P.R.Caron	731
Recombinant DNA data management at the restriction and functional site level D.Shalloway and N.R.Deering	739
The diagonal-traverse homology search algorithm for locating similarities between two sequences C.T.White, S.C.Hardies, C.A.Hutchison, III and M.H.Edgell	751
Analysis of large nucleic acid dot matrices on small computers S.E.Zweig	767

Introduction of restriction enzyme sites in protein-coding DNA sequences by site-specific mutagenesis not affecting the amino acid sequence: a computer program R.Arentzen and W.C.Ripka	777
<i>Escherichia coli</i> promoter sequences predict <i>in vitro</i> RNA polymerase selectivity M.E.Mulligan, D.K.Hawley, R.Entriken and W.R.McClure	789
Graphic display of nucleic acid structure by a microcomputer K.Watanabe, K.Yasukawa and K.Iso	801
Microcomputer programs for graphic analysis of nucleic acid and protein sequences D.W.Mount and B.Conrad	811
Microcomputer programs for back translation of protein to DNA sequences and analysis of ambiguous DNA sequences D.W.Mount and B.Conrad	819
'ZSTATS' — a statistical analysis for potential Z-DNA sequences J.K.Vass and R.H.Wilson	825
Computer programs used to aid in the selection of DNA hybridization probes R.E.Raupach	833
Computer selection of oligonucleotide probes from amino acid sequences for use in gene library screening J.Yang, J.Ye and D.C.Wallace	837
Direct calculation of the sizes of DNA fragments separated by gel electrophoresis using programmes written for a pocket calculator E.J.Gough and N.M.Gough	845
Author index	855

