

See discussions, stats, and author profiles for this publication at:
<https://www.researchgate.net/publication/223282124>

A bibliography on computational molecular biology and genetics

ARTICLE *in* MATHEMATICAL AND COMPUTER MODELLING · JUNE 1992

Impact Factor: 1.41 · DOI: 10.1016/0895-7177(92)90166-I

READS

54

3 AUTHORS, INCLUDING:



Tarynn Witten

Virginia Commonwealth University

122 PUBLICATIONS 1,302 CITATIONS

SEE PROFILE

A BIBLIOGRAPHY ON COMPUTATIONAL MOLECULAR BIOLOGY AND GENETICS

SARAH BARRON, MATTHEW WITTEN, GONGXIAN LIU

Department of Applications Research and Development

University of Texas System—Center for High Performance Computing

Balcones Research Center, CMS 1.154, 10100 Burnet Road

Austin, TX 78758-4497, U.S.A.

Abstract — The field of computational molecular biology and genetics is expanding at an enormous rate. Journals such as CABIOS and Nucleic Acids Research routinely publish articles on computational and mathematical aspects of biology. The purpose of this paper is to provide a bibliographic review of the literature in this area related to DNA mapping and sequence analysis. We have focused on computer and mathematical aspects of molecular biology and genetics (interpreted in a broad sense). Authors are solicited for their additions/corrections to this bibliography. Contact us at the above address.

ACKNOWLEDGEMENTS

We would like to acknowledge the assistance of numerous individuals without whom this bibliography would not have been as complete or as comprehensive. With this in mind, we gratefully acknowledge the help of the following individuals (in no particular order): Amos Bairoch (Department of Medical Biochemistry, University of Geneva, Switzerland), Frederic Fairfield (Center for Human Genome Studies, Los Alamos National Laboratory); Paul Gilna (GenBank, Los Alamos National Laboratory); Robert Jones (Thinking Machines Corporation, Cambridge, MA); Hugh Nicholas (Pittsburg Supercomputing Center); John Wooley (National Science Foundation); Sylvia Spengler (Program in Mathematics and Molecular Biology, UC Berkeley); Susan Naylor (UT Health Science Center, San Antonio); Robert Langridge (Computer Graphics Laboratory, UC San Francisco); Robert Moyzis (Human Genome Project, Los Alamos National Laboratory); George Bell (T-Group, Los Alamos National Laboratory); Creed Abell (College of Pharmacy, University of Texas, Austin); Eldon Sutton (Department of Zoology, University of Texas, Austin); Michael Cinkosky (Los Alamos National Laboratory); Parallel Processing Class (Argonne National Laboratory); Charles DeLisi (College of Engineering, Boston University); S. Pongor (Trieste, Italy); Eugene W. Myers (Departments of Molecular and Cellular Biology and Computer Science, University of Arizona); John Quinn (AMT Digital Array Processor (DAP)); Z. Galil (Weizmann Institute, Israel); Michael S. Waterman (Department of Mathematics and Molecular Biology, University of Southern California); Dave Thurman (Pacific Northwest Laboratories); Pavel Pevzner (Department of Mathematics, University of Southern California); M. S. Gelfand (Institute of Protein Research, USSR Acad. Science, Pushchina, Moscow Region); H. A. Lim (Supercomputer Computations Research Institute, Florida State University); and J. M. Labouygues (GeoGamow Institute, France). All mistakes are the sole responsibility of the authors.

REFERENCES

1. AARONSON, R., YOUNG, J. F., AND PALESE, P. Oligonucleotide mapping: evaluation of its sensitivity by computer-simulation. *Nucl. Acids Res.* **10**(1) (1982), 237-246.
2. ABAGYAN, R. A., AND MAIOROV, V. N. Efficient method for computer-aided search for conformational similarities in database of protein three-dimensional structures, (in) *Abstracts of: Application of Data Bases and Microcomputers in Molecular Biology*, (Jena), p.53, 1988.
3. ABAGYAN, R. A., AND MAIOROV, V. N. A simple qualitative representation of polypeptide chain folds: comparison of protein tertiary structures. *J. Biomol. Struct. Dyn.* **5**(1) (1988), 1267-1279.
4. ABARBANEL, R. M., WIENEKE, P. R., MANSFIELD, E., JAFFE, D. A., AND BRUTLAG, D. L. Rapid searches for complex patterns in biological molecules. *Nucl. Acids Res.* **12** (1984), 263-280.
5. ABOLA, E. E., BERNSTEIN, F. C., BRYANT, S. H., KOETZLE, T. F., AND WENG, J. Protein data bank, (in) *Crystallographic Databases*, (Data Commission of the International Union of Crystallography, Bonn/Cambridge/Chester), pp.107-132, 1987.
6. ABOLA, E. E., BERNSTEIN, F. C., AND KOETZLE, T. F. The protein data bank, (in) *The Role of Data in Scientific Progress*, Glaeser, P.S., (ed.), (Elsevier, North Holland), pp.139-144, 1985.
7. ABOLA, E. E., BERNSTEIN, F. C., AND KOETZLE, T. F. The protein data bank, (in) *Computational Molecular Biology. Sources And Methods For Sequence Analysis*, Lesk, A.M., (ed.), (Oxford University Press, Oxford), pp.69-81, 1988.
8. ABRAHAM, D. J., AND LEO, A. J. Extension of the fragment method to calculate amino acid zwitterion and side chain partition coefficients. *Proteins* **2** (1987), 130-152.
9. ABRAHAMSON, K. Generalized string matching. *SIAM J. Computing* **16** (1987), 1039-1051.
10. ABREMSKI, K., SIROTKIN, K., AND LAPEDES, A. Application of neural networks and information theory to the identification of *E. coli* transcriptional promoters. Tech. Rep. LA-UR-91-729, Los Alamos National Laboratory, 1991.
11. ACALUGARITEI, G., AND PAUN, G. Classification of genetic messages according to the quantitative relations among the types of component nucleotides. *Cybernetica* **32** (1989), 247-262.
12. ADAMS, J., AND ROTHMAN, E. D. Estimations of phylogenetic relationships from DNA restriction patterns and selection of endonuclease cleavage sites. *Proc. Natl. Acad. Sci. USA* **79** (1982), 3560-3564.
13. ADZHUBEI, A. A., EISENMAYER, F., TUMANIAN, V. G., BRODZINSKI, S., AND ESPOVA, N. G. Classification of protein secondary structure types on the basis of protein data bank, (in) *Abstracts of: Application of Data Bases and Microcomputers in Molecular Biology*, (Jena), p.51, 1988.
14. AHO, A. V., AND CORASICK, M. J. Efficient string matching: an aid to bibliographic search. *CACM* **18** (1975), 333-340.
15. AHO, A. V., HIRSCHBERG, D. S., AND ULLMAN, J. D. Bounds on the complexity of the longest common subsequence problem. *J. ACM* **23** (1976), 1-12.
16. AHO, A. V., HOPCROFT, J. E., AND ULLMAN, J. D. *The Design and Analysis of Computer Algorithms*, (Addison-Wesley, Reading, MA), 1974.
17. AKIMOVA, N. I., ALEXANDROV, A. A., NITSAI, O. V., PALMBAKH, G. G., SEDOV, A. E., VITZON, L. A., ZHANGHELDINA, G. T., GOLOVANOV, E. I., SPRIZHITSKY, Y. A., AND SHEPELEV, V. A. GENEXPRESS, database of primary biopolymer structures, (in) *Abstracts of: Application of Data Bases and Microcomputers in Molecular Biology*, (Jena), p.18, 1988.
18. AKRIGG, D., BLEASBY, A. J., DIX, N. I. M., FINDLAY, J. B. C., NORTH, A. C. T., PARRY-SMITH, D., AND WOOTON, J. C. A protein sequence/structure database. *Nature* **335** (1988), 745-746.
19. ALEXANDROV, A. A. The development of databases for biopolymer structures and working procedures in 1985-1988: major results and trends, (in) *Abstracts of: Application of Data Bases and Microcomputers in Molecular Biology*, (Jena), pp.14-15, 1988.
20. ALEXANDROV, N. N., AND MIRONOV, A. A. Pattern recognition in computer analysis of nucleotide sequences. *Mol. Biol.* **23** (1989), 988-999.
21. ALEXANDROV, N. N., AND MIRONOV, A. A. Application of a new method of pattern recognition. *Nucl. Acids Res.* **18** (1990), 1847-1852.
22. ALFF-STEINBERGER, C. Evidence for a coding pattern on the non-coding strand of the *E. coli* genome. *Nucl. Acids Res.* **12** (1984), 2235-2241.
23. ALLEWELL, N. M., AND OBEROI, H. Electrostatic effects in protein folding, stability, and function, (in) *Methods in Enzymology-Molecular design and Modeling: Concepts and Applications*, Part A, Vol.202, Langone, J.J., (ed.), (Academic Press, NY), pp.3-19, 1991.
24. ALLISON, L., AND YEE, C. N. Restriction mapping is in separation theory. *CABIOS* **4** (1988), 97-101.
25. ALLISON, L., AND YEE, C. N. Minimum message length encoding and the comparison of macromolecules. *Bull. Math. Biol.* **52** (1990), 431-453.
26. ALMAGOR, H. A Markov analysis of DNA sequences. *J. Theor. Biol.* **104** (1983), 633-645.
27. ALMAGOR, H. Nucleotide distribution and the recognition of coding regions in DNA sequences: An information theory approach. *J. Theor. Biol.* **117** (1985), 127-136.

28. ALTMAN, R. B. Exclusion Methods for the Determination of Protein Structure from Experimental Data, Ph.D. Thesis, Stanford University Medical Information Sciences, 1989.
29. ALTSCHUL, S. F. Gap costs for multiple sequence alignment. *J. Theor. Biol.* **138** (1989), 297–309.
30. ALTSCHUL, S. F. Leaf pairs and tree dissections. *SIAM J. Disc. Math.* **2**(3) (1989), 293–299.
31. ALTSCHUL, S. F. Amino acid substitution matrices from an information theoretic perspective. *J. Mol. Biol.* **219** (1991), 555–565.
32. ALTSCHUL, S. F., CARROLL, R. J., AND LIPMAN, D. J. Weights for data related by a tree. *J. Mol. Biol.* **207** (1989), 647–653.
33. ALTSCHUL, S. F., AND ERICKSON, B. W. Recognition of good alignments of segments from two DNA or RNA. *Fed. Proc.* **44** (1985), 1086.
34. ALTSCHUL, S. F., AND ERICKSON, B. W. Significance for nucleotide sequence alignments: a method for random sequence permutation that preserves dinucleotide and codon usage. *Mol. Biol. Evol.* **2** (1985), 526–538.
35. ALTSCHUL, S. F., AND ERICKSON, B. W. Locally optimal subalignments using nonlinear similarity functions. *Bull. Math. Biol.* **48** (1986), 633–660.
36. ALTSCHUL, S. F., AND ERICKSON, B. W. A nonlinear measure of subalignment similarity and its significance levels. *Bull. Math. Biol.* **48** (1986), 617–632.
37. ALTSCHUL, S. F., AND ERICKSON, B. W. Optimal sequence alignment using affine gap costs. *Bull. Math. Biol.* **48** (1986), 603–616.
38. ALTSCHUL, S. F., AND ERICKSON, B. W. Significance levels for biological sequence comparison using non-linear similarity functions. *Bull. Math. Biol.* **50** (1988), 77–92.
39. ALTSCHUL, S. F., GISH, W., MILLER, W., MYERS, E. W., AND LIPMAN, D. J. Basic local alignment search tool. *J. Mol. Biol.* **215** (1990), 403–410.
40. ALTSCHUL, S. F., AND LIPMAN, D. J. Trees, stars, and multiple biological sequence alignment. *SIAM J. Appl. Math.* **49** (1989), 197–209.
41. ALTSCHUL, S. F., AND LIPMAN, D. J. Protein database searches for multiple alignments. *Proc. Natl. Acad. Sci.* **87** (1990), 5509–5513.
42. AMIR, A., AND LANDAU, G. M. Fast parallel and serial multidimensional approximate array matching, (in) *Sequences, combinatorics, compression, security and transmission*, Capocelli, R.M., (Springer-Verlag), pp.3–24, 1990.
43. ANDERSON, G. C. Genome database booms as journals take the hard line. *The Scientist*, October 30 (1989), 4–5.
44. ANDERSON, G. C. Creation of linkage map falters, posing delay for genome project. *The Scientist*, January 8 (1990).
45. AOE, J. An efficient implementation of static string pattern matching machines. *IEEE Transactions on Software Engineering* **15** (1989), 1010–1016.
46. AOTA, S., AND IKEMURA, T. Diversity in G+C content of the third position of codons in vertebrate genes and its cause. *Nucl. Acids Res.* **14** (1986), 6345–6356.
47. AOTA, S-I. GOJOBORI, T., ISHIBASHI, F., MARUYAMA, T., AND IKEMURA, T. Codon usage tabulated from the GenBank genetic sequence data. *Nucl. Acids Res.* **16** (1988), r315–r323.
48. APOSTOLICO, A. Parallel log-time construction of suffix trees. Tech. Rep. CSD-TR-632, Purdue, 1986.
49. APOSTOLICO, A., AND PREPARATA, F. P. Structural properties of the string statistics problem. *J. Computer and System Sciences* **31** (1985), 394–411.
50. ARENTZEN, R., AND RIPKA, W. C. Introduction to restriction enzyme sites in protein-coding DNA sequences by site-specific mutagenesis not affecting amino acid sequences: a computer program. *Nucl. Acids Res.* **12** (1984), 777–787.
51. ARGOS, P. Analysis of sequence-similar pentapeptides in unrelated protein tertiary structures: strategies for protein folding and a guide for site-directed mutagenesis. *J. Mol. Biol.* **197** (1987), 331–348.
52. ARGOS, P. A sensitive procedure to compare amino acid sequences. *J. Mol. Biol.* **193** (1987), 385–396.
53. ARGOS, P., HANEI, M., AND GARAVITO, R. M. The Chou-Fasman secondary structure prediction method with an extended data base. *FEBS Lett.* **93** (1978), 19–24.
54. ARGOS, P., HANEI, M., WILSON, J., AND KELLEY, W. N. A possible nucleotide-binding domain in the tertiary fold of phosphoribosyl transferases. *J. Biol. Chem.* **258** (1983), 6450–6457.
55. ARGOS, P., AND McCALDON, P. Theoretical and computer analysis of protein primary sequences: Structure comparison and prediction, (in) *Genetic Engineering, Principle and Methods*, Setlow, J.K., (ed.), 10:21–66, (Plenum Press, NY), 1988.
56. ARGOS, P., AND PALAU, J. Amino acid distribution in protein secondary structures. *Int. J. Pept. Protein Res.* **19** (1982), 380–393.
57. ARGOS, P., AND RAO, J. K. M. Prediction of protein structure, (in) *Methods in Enzymology*, Vol.130, (Academic Press, NY), pp.185–207, 1986.
58. ARGOS, P., RAO, J. K. M., AND HARGRAVE, P. A. Structural prediction of membrane-bound proteins. *Eur. J. Biochem.* **128** (1982), 565–575.
59. ARGOS, P., ROSSMAN, M. G., GRAU, U. M., ZUBER, H., FRANK, G., AND TRATSCHIN, J. D. Thermal stability and protein structure. *Biochemistry* **18** (1979), 5698–5703.

60. ARGOS, P., AND SCHWARZ, J. An assessment of protein secondary structure prediction methods based on amino acid sequence. *Biochim. Biophys. Acta* **439** (1976), 261-273.
61. ARGOS, P., AND SIEZEN, R. J. Structural homology of lens crystallins. A method to detect protein structural homology from primary sequences. *Eur. J. Biochem.* **131** (1983), 143-148.
62. ARGOS, P., AND VINGRON, M. Sensitivity comparison of protein amino acid sequences, (in) **Methods in Enzymology-Molecular Evolution: Computer Analysis of Protein and Nucleic Acid Sequences**, Vol.183, Doolittle, R.F., (ed.), (Academic Press, NY), pp.352-365, 1990.
63. ARNOLD, J., CUTICCHIA, A. J., NEWSOME, D. A., JENNINGS III, W. W., AND IVARIE, R. Mono-through hexanucleotide composition of the sense strand of yeast DNA: a Markov chain analysis. *Nucl. Acids Res.* **16** (1988), 7145-7158.
64. ARQUES, D. G., AND MICHEL, C. J. Periodicities in introns. *Nucl. Acids Res.* **15** (1987), 7581-7592.
65. ARQUES, D. G., AND MICHEL, C. J. A purine-pyrimidine motif verifying an identical presence in almost all taxonomic groups. *J. Theor. Biol.* **128** (1987), 457-461.
66. ARQUES, D. G., AND MICHEL, C. J. Study of a perturbation in the coding periodicity. *Math. Biosci.* **86** (1987), 1-14.
67. ARQUES, D. G., AND MICHEL, C. J. A model of DNA sequence evolution. *Bull. Math. Biol.* **52** (1990), 741-772.
68. ARQUES, D. G., AND MICHEL, C. J. Periodicities in coding and noncoding regions of the genes. *J. Theor. Biol.* **143** (1990), 307-318.
69. ARQUES, D. G., MICHEL, C. J., AND ORIEUX, K. Analysis of gene evolution: the software AGE. *CABIOS* **8(1)** (1992), 5-14.
70. ARRATIA, R., AND GORDON, L. Tutorial on large deviations for the binomial distribution. *Bull. Math. Biol.* **51(1)** (1989), 125-131.
71. ARRATIA, R., GORDON, L., AND WATERMAN, M. An extreme value theory for sequence matching. *Ann. Stat.* **14** (1986), 971-993.
72. ARRATIA, R., MORRIS, P., AND WATERMAN, M. Stochastic scrabble: large deviations for sequences with scores. *J. Appl. Prob.* **25** (1988), 106-119.
73. ARRATIA, R., AND WATERMAN, M. An Erdos-Renyi law with shifts. *Adv. Math.* **55** (1985), 13-23.
74. ARRATIA, R., AND WATERMAN, M. Critical phenomena in sequence matching. *Ann. Prob.* **13** (1985), 1236-1249.
75. ARRATIA, R., AND WATERMAN, M. The Erdos-Renyi strong law for pattern matching with a given proportion of mismatches. *Ann. Prob.* **17** (1989), 1152-1169.
76. ARRIGO, P., GIULIANO, F., SCALIA, F., RAPALLO, A., AND DAMIANI, G. Identification of a new motif on nucleic acid sequence data using Kohonen's self-organizing map. *CABIOS* **7(3)** (1991), 353-357.
77. ARTECA, G. A., AND MEZEY, P. G. A method for the characterization of foldings in protein ribbon models. *J. Mol. Graphics* **8** (1990), 66-80.
78. ARTECA, G. A., TAPIA, O., AND MEZEY, P. G. Implementing knot-theoretical characterization methods to analyze the backbone structure of proteins: Application to CTF L7/L12 and carboxypeptidase A inhibitor proteins. *J. Mol. Graphics* **9(3)** (1991), 148-156.
79. ARTYMIUK, P. J., RICE, D. W., MITCHELL, E. M., AND WILLETT, P. Searching techniques for databases of protein secondary structures. *J. Info. Sci.* **15** (1989), 287-298.
80. ATILGAN, T., NICHOLAS, H. B. Jr., AND MCCLAIN, W. H. A statistical method for correlating tRNA sequence with amino acid specificity. *Nucl. Acids Res.* **1** (1986), 375-380.
81. ATLAN, H., AND KOPPEL, M. The cellular computer DNA: program or data. *Bull. Math. Biol.* **52** (1990), 335-348.
82. ATTIMONELLI, M., LANAVE, C., SBISA, E., PREPARATA, G., AND SACCOME, C. Multisequence comparisons in protein coding genes. *Cell Biophysics* **7** (1985), 239-250.
83. AUGER, I. E., AND LAWRENCE, C. E. Algorithms for the optimal identification of segment neighborhoods. *Bull. Math. Biol.* **51(1)** (1989), 39-54.
84. AURON, P. E., RINDONE, W. P. C., VARY, P. H., CELENTANO, J. J., AND VOURLAKIS, J. N. Computer-aided prediction of RNA secondary structures. *Nucl. Acids Res.* **10(1)** (1982), 403-419.
85. AVERY, P. J. The analysis of intron data and its use in the detection of short signals, *J. Mol. Evol.*, (submitted).
86. AYER, D., AND YARUS, M. The context effect does not require a fourth base pair. *Science* **231** (1986), 393-395.
87. BABENKO, V. E., AND SAGITOV, V. R. The vectors data base, (in) **Abstracts of: Application of Data Bases and Microcomputers in Molecular Biology**, (Jena), p.26, 1988.
88. BACH, R., FRIEDLAND, P., BRUTLAG, D. L., AND LEDES, L. MAXAMIZE: A DNA sequencing strategy advisor. *Nucl. Acids Res.* **10** (1982), 295-304.
89. BACH, R., IWASAKI, Y., AND FRIEDLAND, P. Intelligent computational assistance for experiment design. *Nucl. Acids Res.* **12** (1984), 11-29.
90. BACHMANN, B. J. Linkage map of *Escherichia coli* K-12, Edition 8. *Microbiol. Rev.* **54** (1990), 130-197.
91. BACON, D. J., AND ANDERSON, W. F. Multiple sequence alignment. *J. Mol. Biol.* **191** (1986), 153-161.

92. BACON, D. J., AND ANDERSON, W. F. Multiple sequence comparison, (in) **Methods in Enzymology-Molecular Evolution: Computer Analysis of Protein and Nucleic Acid Sequences**, Vol.183, Doolittle, R.F., (ed.), (Academic Press, NY), pp.438-447, 1990.
93. BAEZA-YATES, R. A. Improved string searching. *Software Practice and Experience* 19 (1989), 257-271.
94. BAINS, W. MULTAN (2), a multiple string alignment program for nucleic acids and proteins. *CABIOS* 5(1) (1989), 51-52.
95. BAINS, W., AND SMITH, G. C. A novel method for nucleic acid sequence determination. *J. Theor. Biol.* 135 (1988), 303-307.
96. BAIROCH, A. SEQANALREF: a sequence analysis bibliographic reference databank. *CABIOS* 7(2) (1991), 268.
97. BALABAN, A. T. Highly discriminating distance-based topological index. *Chemical Physics Letters* 89(5) (1982), 399-403.
98. BALDING, D. J., AND TORNEY, D. C. Statistical analysis of DNA fingerprint data for ordered clone physical mapping of human chromosomes. *Bull. Math. Biol.* 53(6) (1991), 853-879.
99. BANDELT, H. J., AND DRESS, A. W. M. Weak hierarchies associated with similarity measure-an additive clustering technique. *Bull. Math. Biol.* 51(1) (1989), 133-166.
100. BARBER, A. M., AND MAIZEL, J. V. Sequence Editing Aligner: A multiple sequence editor and aligner. *Gene Anal. Techn. Appl.* 7 (1990), 39-45.
101. BARBER, A. M., AND ZHURKIN, V. B. CAP binding sites reveal pyrimidine-purine pattern characteristic of DNA bending. *J. Biomol. Struct. Dyn.* 8 (1990), 213-232.
102. BARCISZEWSKI, J., JOACHIMIAK, A., AND TWAROWSKI, T. Transfer and ribosomal 5S ribonucleic acids data bases, (in) **Abstracts of: Application of Data Bases and Microcomputers in Molecular Biology**, (Jena), p.42, 1988.
103. BARKER, W. C., AND DAYHOFF, M. O. Detecting distant relationships: computer methods and results, (in) **Atlas of Protein Sequence and Structure**, 5:101-110, (Nat. Biomed. Res. Found., Washington, D.C.), 1972.
104. BARKER, W. C., GEORGE, D. G., AND HUNT, L. T. Protein sequence database, (in) **Methods in Enzymology-Molecular Evolution: Computer Analysis of Protein and Nucleic Acid Sequences**, Vol.183, Doolittle, R.F., (ed.), (Academic Press, NY), pp.31-49, 1990.
105. BARKER, W. C., HUNT, L. T., AND GEORGE, D. G. Research uses of protein sequence databases. *Protein, Nucleic Acid and Enzyme* 29 (1986), 163-176.
106. BARKER, W. C., HUNT, L. T., AND GEORGE, D. G. Identifying domains in protein sequences. *Protein Seq. Data Anal.* 1 (1988), 363-373.
107. BARKER, W. C., HUNT, L. T., GEORGE, D. G., AND ORCUTT, B. C. A resource for protein identification, (in) **The Role of Data in Scientific Progress**, Glaeser, P.S., (ed.), (Elsevier, North Holland), pp.127-133, 1985.
108. BARKER, W. C., HUNT, L. T., SEIBEL-ROSS, E., YEH, L.-S., AND GEORGE, D. G. Identifying protein domains using consensus sequences. *Fed. Proc.* 46 (1987), 2232.
109. BARKER, W. C., JOHNSON, G. C., HUNT, L. T., AND GEORGE, D. G. Similar domains in different proteins: detection and significance. *Protein, Nucleic Acid and Enzyme* 29 (1986), 54-68.
110. BARKER, W. C., KETCHAM, L. K., AND DAYHOFF, M. O. A comprehensive examination of protein sequences for evidence of internal gene duplication. *J. Mol. Evol.* 10 (1978), 265-281.
111. BARKOWSKY, E. V. Prediction of the secondary structure of globular proteins by their amino acid sequence. *Acta Biol. Med. Ger.* 41 (1982), 751-758.
112. BARNETT, L. J. Probabilistic analysis of random clone restriction mapping. Master's thesis, Washington University in St. Louis, 1990.
113. BARNHART, B. J. The Department of Energy (DOE) human genome initiative. *Genomics* 5 (1989), 657-660.
114. BARRAI, I., SCAPOLI, C., BARALE, R., AND VOLINIA, S. Oligonucleotide correlations between infector and host genomes hint at evolutionary relationships. *Nucl. Acids Res.* 18 (1990), 3021-3025.
115. BARRON, S., WITTEN, M., HARKNESS, R., AND DRIVER, J. A bibliography on computational algorithms in molecular biology and genetics. *CABIOS* 7(2) (1991), 269.
116. BARRON, S., WITTEN, M., HARKNESS, R., WANG, F., AND DRIVER, J. The Human Genome Project: An Overview of Computational Issues in Molecular Biology and Genetics, (in) **Alternatives to Animal Testing**, in press.
117. BARRY, D., AND HARTIGAN, J. A. Asynchronous distance between homologous DNA sequences. *Biometrics* 43 (1987), 261-276.
118. BARRY, D., AND HARTIGAN, J. A. Statistical analysis of hominid molecular evolution. *Stat. Sci.* 2 (1987), 191-210.
119. BARSALOU, T., AND BRUTLAG, D. L. Searching gene and protein sequence databases. *M. D. Computing* 8(3) (1991), 144-149.
120. BARTON, G. J. Protein multiple sequence alignment and flexible pattern matching, (in) **Methods in Enzymology-Molecular Evolution: Computer Analysis of Protein and Nucleic Acid Sequences**, Vol.183, Doolittle, R.F., (ed.), (Academic Press, NY), pp.403-428. 1990.

121. BARTON, G. J. Scanning protein sequence databanks using a distributed processing workstation network. *CABIOS* 7(1) (1991), 85-88.
122. BARTON, G. J., AND STERNBERG, M. J. E. Evaluation and improvements in the automatic alignment of protein sequences. *Protein Engineering* 1 (1987), 89-94.
123. BARTON, G. J., AND STERNBERG, M. J. E. A strategy for the rapid multiple alignment of protein sequences. *J. Mol. Biol.* 198 (1987), 327-337.
124. BARTON, G. J., AND STERNBERG, M. J. E. LOPAL and SCAMP: techniques for the comparison and display of protein structures. *J. Mol. Graphics* 6 (1988), 190-196, 200-201.
125. BARTON, G. J., AND STERNBERG, M. J. E. Flexible protein sequence patterns: a sensitive method to detect weak structural similarities. *J. Mol. Biol.* 212 (1990), 389-402.
126. BARTON, N. H., AND JOHNES, J. S. The language of the genes. *Nature* 346 (1990), 415-416.
127. BARYSZ, M., JASHARI, G., LALL, R. S., SRIVASTAVA, V. K., AND TRINAJSTIC, N. On the distance matrix of molecules containing heteroatoms, (in) *Chemical Applications of Topology and Graph Theory*, King, R.B., (ed.), (Elsevier, Amsterdam), 1983.
128. BASH, P. A., PATTABIRAMAN, N., HUANG, C., FERRIN, T. E., AND LANGRIDGE, R. Van der Waals surfaces in molecular modeling: Implementation with real-time computer graphics. *Science* 222 (1983), 1325-1327.
129. BASTIE-SIGEAC, F., BELLIS, M., AND ROIZES, G. A method for selecting the coding DNA for a protein of known sequence. II. Use of obligatory restriction sites. *Biochimie* 65 (1983), 77-82.
130. BATCHINSKY, A. G., AND RATNER, V. A. Noise immunity of the Genetic Code. *Biometr. Zeitschrift* 18 (1976), 53.
131. BECKMANN, J. S., BRENDL, V., AND TRIFONOV, E. N. Intervening sequences exhibit distinct vocabulary. *J. Biomol. Struct. Dyn.* 4(3) (1986), 391-400.
132. BELL, G. I., AND MARR, T. G. , (eds.). *Computers and DNA*, Proceedings of the Interface Between Computation Science and Nucleic Acid Sequencing Workshop, Vol.VII, (Addison-Wesley, Reading, Massachusetts), 1990.
133. BELLON, B. Construction of restriction maps. *CABIOS* 4 (1988), 111-116.
134. BENEDETTI, G., De Santis, P., AND MOROSETTI, S. A new method to find a set of energetically optimal RNA secondary structures. *Nucl. Acids Res.* 17 (1989), 5249-5161.
135. BENEDETTI, G., AND MOROSETTI, S. Three-dimensional folding of *Tetrahymena thermophila* rRNA IVS sequence: A proposal. *J. Biomol. Struct. Dyn.* 8(5) (1991), 1045-1055.
136. BENGIO, Y., AND POULIOT, Y. Efficient recognition of immunoglobulin domains from amino acid sequences using a neural network. *CABIOS* 6(4) (1990), 319-324.
137. BENHAM, C. J. Geometry and mechanics of DNA superhelicity. *Biopolymers* 22 (1983), 2477-2495.
138. BENHAM, C. J. Theoretical analysis of conformational equilibria in superhelical DNA. *Ann. Rev. Biophys. Biophys. Chem.* 14 (1985), 23-51.
139. BENHAM, C. J. Mechanics and equilibria of superhelical DNA, (in) *Mathematical Methods For DNA Sequences*, Waterman, M.S., (ed.), (CRC Press, Boca Raton, FL), 1989.
140. BENHAM, C. J. Theoretical analysis of heteropolymeric transitions in superhelical DNA molecules of specified sequence. *J. Chem. Phys.* 92(10) (1990), 6294-6305.
141. BENNETZEN, J. L., AND HALL, B. D. Codon selection in yeast. *J. Biol. Chem.* 257 (1982), 3026-3031.
142. BENSON, D., BOGUSKI, M., LIPMAN, D. J., AND OSTELL, J. The National Center for Biotechnology Information. *Genomics* 6 (1990), 389-391.
143. BENSON, D. C. Digital signal processing methods for biological sequence comparison. *Nucl. Acids Res.* 18 (1990), 3001-3006.
144. BENSON, D. C. Fourier methods for biosequence analysis. *Nucl. Acids Res.* 18 (1990), 6305-6310.
145. BENTLEY, J. L. Tools for Experiments on Algorithms, Bell Labs, (Murray Hill, NJ), 1990.
146. BENZER, S. On the topology of the genetic fine structure. *Proc. Natl. Acad. Sci.* 45 (1959), 1607-1620.
147. BENZER, S. The fine structure of the gene. *Scientific American* 206 (1962), 70-84.
148. BERAN-KOEHN, J. C., AND GILLETT, W. D. Information theoretic estimation of clone overlap probabilities. Tech. Rep. WUCS-90-27, Washington University in St. Louis, 1990.
149. BERG, O. G. Selection of DNA binding sites by regulatory proteins: Functional specificity and pseudosite competition. *J. Biomol. Struct. Dyn.* 6(2) (1988), 275-297.
150. BERG, O. G., AND VON HIPPEL, P. H. Selection of DNA binding sites by regulatory proteins. I. Statistical-mechanical theory and application to operators and promoters. *J. Mol. Biol.* 193 (1987), 723-750.
151. BERG, O. G., AND VON HIPPEL, P. H. Selection of DNA binding sites by regulatory proteins. *Trends Biochem. Sci.* 13 (1988), 207-211.
152. BERG, O. G., AND VON HIPPEL, P. H. Selection of DNA binding sites by regulatory proteins. II. The binding specificity of cyclic AMP receptor to recognition sites. *J. Mol. Biol.* 200 (1988), 709-723.
153. BERGER, M. P., AND MUNSON, P. J. A novel randomized iterative strategy for aligning multiple protein sequences. *CABIOS* 7(4) (1991), 479-484.
154. BERKMAN, O., JAJA, J., KRISHNAMURTHY, S., THURIMELLA, R., AND VISHKIN, U. Some triply-logarithmic parallel algorithms, (in) *IEEE Symposium on the Foundations of Computer Science*, 1990.

155. BERNARDI, G., AND BERNARDI, G. Codon usage and genome composition. *J. Mol. Evol.* **22** (1985), 363-365.
156. BERNARDI, G., AND BERNARDI, G. Compositional constraints and genome evolution. *J. Mol. Evol.* **24** (1986), 1-11.
157. BERNARDI, G., AND BERNARDI, G. Compositional patterns in the nuclear genomes of cold-blooded vertebrates. *J. Mol. Evol.* **31** (1990), 265-281.
158. BERNARDI, G., MOUCHIROUD, D., GAUTIER, C., AND BERNARDI, G. Compositional patterns in vertebrate genomes: conservation and change in evolution. *J. Mol. Evol.* **28** (1988), 7-18.
159. BERNARDI, G., OLOFFSON, B., FILIPSKI, J., ZERIAL, M., SALINAS, J., GUNY, G., MEUNIER-ROTIVAL, M., AND RODIER, F. The mosaic genome of warm-blooded vertebrates. *Science* **228** (1985), 953-958.
160. BERNHARD, S. A., BRADLEY, D. F., AND DUDA, M. L. Automatic determination of amino acid sequences. *IBM J. Res. Dev.* **7** (1963), 246-251.
161. BERNSTEIN, F. C., KOETZLE, T. F., WILLIAMS, G. J. B., MEYER, E. F., BRICE, M. D., RODGERS, J. R., KENNARD, O., SHIMANOUCHI, T., AND TASUMI, M. The protein data bank: A computer-based archival file for macromolecular structures. *J. Mol. Biol.* **112** (1977), 535-542.
162. BERNSTEIN, M. Reducing the man - machine barrier: the sequence analysis workbench. *CABIOS* **3** (1987), 229-232.
163. BERNSTEIN, M. Using spreadsheet languages to understand sequence analysis algorithms. *CABIOS* **3** (1987), 217-221.
164. BERTOSSI, A. A., LODI, E., LUCCIO, F., AND PAGLI, L. Context-dependent string matching, (in) *Sequences, Combinatorics, Compression, Security and Transmission*, Capocelli, R.M., (Springer-Verlag), p.25-40, 1990.
165. BERTZ, S. H. A mathematical model of molecular complexity, (in) *Chemical Applications of Topology and Graph Theory*, King, R.B., (ed.), (Elsevier/North-Holland, Amsterdam), 1983.
166. BEUTEL, B. A., AND RECORD, M. T. Jr. *E. coli* promoter spacer regions contain nonrandom sequences which correlate to spacer length. *Nucl. Acids Res.* **18** (1990), 3597-3603.
167. BEUTLER, E., GELBART, T., HAN, J., KOZIOL, J., AND BEUTLER, B. Evolution of the genome and the genetic code: selection at the dinucleotide level by methylation and polyribonucleotide cleavage. *Proc. Natl. Acad. Sci.* **86** (1989), 192-196.
168. BEVERLEY, S. M., AND C., W. A. Molecular evolution in and the higher Diptera II. A time scale for fly evolution. *J. Mol. Evol.* **21** (1984), 1-13.
169. BEYER, W. A., STEIN, M. L., SMITH, T. F., AND ULAM, S. M. A molecular sequence metric and evolutionary trees. *Math. Biosci.* **19** (1974), 9-25.
170. BHASKARAN, R., AND PONNUSWAMY, P. K. Positional flexibilities of amino acid residues in globular proteins. *Int. J. Pept. Protein. Res.* **32** (1988), 242-255.
171. BIBB, M. J., FINDLAY, P. R., AND JOHNSTON, M. W. The relationship between base composition and codon usage in bacterial genes and its use for the simple and reliable identification of protein-coding sequences. *Gene* **30** (1984), 157-166.
172. BICKNELL, E. J., RADA, R., DAVIDSON, S., AND STANDER, R. Mapping from GenBank to MEDLINE. *Nucl. Acids Res.* **16** (1988), 1667-1680.
173. BILARDI, G., AND PREPARATA, F. P. Size-time complexity of boolean networks for prefix computations. *J. ACM* **36** (1989), 362-382.
174. BILOFSKY, H. S., AND BURKS, C. The GenBank genetic sequence data bank. *Nucl. Acids Res.* **16** (1988), 1861-1863.
175. BILOFSKY, H. S., BURKS, C., FICKETT, J. W., GOAD, W. B., LEWITTER, F. I., RINDONE, W. P., SWINDELL, C. D., AND TUNG, C.-S. The GenBank genetic sequence databank. *Nucl. Acids Res.* **14** (1986), 1-4.
176. BIOU, V., GIBRAT, J. F., LEVIN, J. M., ROBSON, B., AND GARNIER, J. Secondary structure prediction: combination of three different methods. *Protein Engineering* **2**(9) (1988), 185-191.
177. BIRD, A. P. DNA methylation and the frequency of CpG in animal DNA. *Nucl. Acids Res.* **8** (1980), 1499-1504.
178. BIRD, A. P. CpG-rich islands and the function of DNA methylation. *Nature* **321** (1986), 209-213.
179. BISHOP, D. T., WILLIAMSON, J. A., AND SKOLNICK, M. H. A model for restriction: Fragment length distribution. *Am. J. Human Genetics* **35** (1983), 795-815.
180. BISHOP, M. J., AND FRIDAY, A. E. Evolutionary trees from nucleic acid and protein sequences. *Proc. R. Soc. Lond. B* **226** (1985), 271-302.
181. BISHOP, M. J., GINSBURG, M., RAWLINGS, C., AND WAKEFORD, R. Molecular sequence databases, (in) *Nucleic Acid and Protein Sequence Analysis: A Practical Approach*, Bishop, M.J., Rawlings, C.J., (eds.), (IRL Press, Oxford), pp.83-113, 1987.
182. BISHOP, M. J., AND RAWLINGS, C. J., (eds.). *Nucleic Acid And Protein Sequence Analysis: A Practical Approach*, (IRL Press, Oxford), 1987.
183. BISHOP, M. J., AND THOMPSON, E. A. Fast computer search for similar DNA sequences. *Nucl. Acids Res.* **12** (1984), 5471-5475.

184. BISHOP, M. J., AND THOMPSON, E. A. Maximum likelihood alignment of DNA sequences. *J. Mol. Biol.* **190** (1986), 159–165.
185. BITTON, D., DEWITT, D. J., HSIAO, D. K., AND MENON, J. A taxonomy of parallel sorting. *Computing Surveys* **16** (1984), 287.
186. BLAISDELL, B. E. Choice of base at silent codon site 3 is selectively neutral in eucaryotic structural genes: it maintains excess short runs of weak and strong hydrogen bonding bases. *J. Mol. Evol.* **19** (1983), 226–236.
187. BLAISDELL, B. E. A prevalent persistent global nonrandomness that distinguishes coding and non-coding eukaryotic nuclear DNA sequences. *J. Mol. Evol.* **19** (1983), 122–133.
188. BLAISDELL, B. E. Markov chain analysis finds a significant influence of neighboring bases on the occurrence of a base in eucaryotic nuclear DNA sequences both protein-coding and noncoding. *J. Mol. Evol.* **21** (1985), 278–288.
189. BLAISDELL, B. E. A measure of the similarity of sets of sequences not requiring sequence alignment. *Proc. Natl. Acad. Sci USA* **83** (1986), 5155–5159.
190. BLAKE, R. D., AND EARLEY, S. Distribution and evolution of sequence characteristics in the *E. coli* genome. *J. Biomol. Struct. Dyn.* **4** (1986), 291–307.
191. BLANGERO, J., KAMMERER, C., AND MACLUEER, J. Multivariate quantitative genetic analysis of genotype-environment interaction. *Am. J. Human Genetics* **43** (1988), A211.
192. BLANKEN, R. L., KLOTZ, L. C., AND HINNEBUSCH, A. G. Computer comparison of new and existing criteria for constructing evolutionary trees from sequence data. *J. Mol. Evol.* **19** (1982), 9–19.
193. BLEASBY, A. J., AND WOOTTON, J. C. Construction of validated, non-redundant composite protein sequence databases. *Protein Engineering* **3** (1990), 153–159.
194. BLOCH, D., BONCHEV, D., BALABAN, A. T., AND MEKENYAN, O. Generalization of the graph center concept and derived topological centric index. *J. Chem. Inf. Comput. Sci.* **20** (1980), 106–113.
195. BLOCH, D. P., MCARTHUR, B., WIDDOWSON, R., SPECTOR, D., GUIMARAES, R. C., AND SMITH, J. tRNA-rRNA sequence homologies: evidence for a common evolutionary origin? *J. Mol. Evol.* **19** (1983), 420–428.
196. BLOSE, S. H., AND HAMBURGER, S. A. Computer-analyzed high resolution two-dimensional gel electrophoresis: a new window for protein research. *Biotechniques* **3** (1985), 232–236.
197. BLUM, A., JIANG, T., LI, M., TROMP, J., AND YANNAKAKIS, M. Linear approximation of shortest superstrings, (in) *Proceedings of the ACM Symposium on Theory of Computing*, ACM press, (Baltimore, MD), 1991.
198. BLUM, R. L. Discovery and representation of causal relationships from a large time-oriented clinical database: The RX project, Ph.D. Thesis, Stanford University. *Computers and Biomedical Research* **15** (1982), 164–187.
199. BLUNDELL, L., COOPER, J., DONNELLY, D., DRIESSEN, H., EDWARDS, Y., EISENMAYER, F., FRAZAO, C., JOHNSON, M., NEIFIND, K., NEWMAN, M., OVERINGTON, J., SALI, A., SLINGSBY, C., AND NALINI, V. ZHU, Z. Patterns of sequence variation in families of homologous proteins, (in) *Methods in Protein Sequence Analysis*, Jornvall, H., Hoog, J.-O., and Gustavsson, A.-M., (ed.), (Birkhauser Verlag, Boston, MA), pp.373–385, 1991.
200. BLUNDELL, T. L., CARNEY, D., GARDNER, S., HAYES, F., HOWLIN, B., HUBBARD, T. J. P., OVERINGTON, J., SINGH, D., SIBANDA, B., AND SUTCLIFFE, M. Knowledge-based protein modelling and design. *Eur. J. Biochem.* **172** (1988), 513–520.
201. BLUNDELL, T. L., SIBANDA, B. L., STERNBERG, M. J. E., AND THORNTON, J. M. Knowledge-based prediction of protein structures and the design of novel molecules. *Nature* **326** (1987), 347–352.
202. BLUNDELL, T. L., AND STERNBERG, M. J. E. Computer-aided design in protein engineering. *Trends Biotech.* **3** (1985), 228–235.
203. BOBROW, D. G., AND HAYES, P. J. , (eds). *AI Journal Special Issue on Qualitative Reasoning about Physical Systems*, (North-Holland), 1984.
204. BODNAR, J. W., AND WARD, D. C. Highly recurring sequence elements identified in eukaryotic DNAs by computer analysis are often homologous to regulatory sequences or protein binding sites. *Nucl. Acids Res.* **15** (1987), 1835–1851.
205. BOGUSKI, M. S., FREEMAN, M., ELSOURBAGY, N., TAYLOR, J., AND GORDON, J. I. On computer-assisted analysis of biological sequences: proline punctuation, consensus sequences and apolipoprotein repeats. *J. Lip. Res.* **27** (1986), 1011–1034.
206. BOHR, H., BOHR, J., BRUNAK, S., COTTERILL, R. M. J., LAUTRUP, B., NORSKOV, L., OLSEN, O., AND PETERSEN, S. B. Protein secondary structure and homology by neural networks. *FEBS Lett.* **241** (1988), 223–228.
207. BONHAM-SMITH, P. C., AND BOURQUE, D. P. Translation of chloroplast-encoded mRNA: potential initiation and termination signals. *Nucl. Acids Res.* **17** (1989), 2057–2080.
208. BOOTH, K. S., AND LUEKER, G. S. Testing for the consecutive ones property, interval graphs, and graph planarity using PQ-tree algorithms. *J. Computer and System Sciences* **13** (1976), 335–379.
209. BORK, P., AND GRUNWALD, C. Pattern search in protein sequence data bases: detection of different nucleotides-binding sites, (in) *Abstracts of: Application of Data Bases and Microcomputers in Molecular Biology*, (Jena), p.45, 1988.

210. BORODOVSKY, M. Y., AND GUSEIN-ZADE, S. M. A general rule for ranged series of codon frequencies in different genomes. *J. Biomol. Struct. Dyn.* 6 (1989), 1001-1012.
211. BOSSI, L., AND ROTH, J. The influence of codon context in genetic code translation. *Nature* 286 (1980), 123-127.
212. BOSSINGER, J. A. The Annotator's Assistant: an expert system for direct submission of genetic sequence data. *CABIOS* 4 (1988), 197-202.
213. BOSWELL, D. R. Sequence alignment by word processor. *Trends Biochem. Sci.* 12 (1987), 279-280.
214. BOSWELL, D. R., AND LESK, A. M. Sequence comparison and alignment: The measurement and interpretation of sequence similarity, (in) *Computational Molecular Biology: Sources and Methods for Sequence Analysis*, Lesk, A.M., (ed.), (Oxford University Press, NY), 1988.
215. BOSWELL, D. R., AND McLACHLAN, A. D. Sequences comparison by exponentially-damped alignment. *Nucl. Acids Res.* 12 (1984), 457-464.
216. BOUGUERET, L., TEKAIA, F., SAUVAGET, I., AND CLAVERIE, J.-M. Objective comparison of exon and intron sequences by means of 2-dimensional data analysis methods. *Nucl. Acids Res.* 16 (1988), 1729-1738.
217. BOYER, R. S., AND MOORE, J. S. A fast string searching algorithm. *CACM* 20 (1977), 762-772.
218. BRADFORD, J. H. Sequence matching with binary codes. *Info. Proc. Letters* 34 (1990), 193-196.
219. BRANSCOMB, E., SLEZAK, T., PAE, R., GALAS, D., CARRANO, A. V., AND WATERMAN, M. Optimizing restriction fragment fingerprinting methods for ordering large genomic libraries. *Genomics* 8 (1990), 351-366.
220. BRAUN, W., BOSCH, C., BROWN, L. R., GO, N., AND WUTHRICH, K. Combined use of proton-proton overhauser enhancements and a distance geometry algorithm for determination of polypeptide conformations. Application to micelle-bound glucagon. *Biochim. Biophys. Acta* 667 (1981), 377-396.
221. BREEN, S., WATERMAN, M. S., AND ZHANG, N. Renewal theory for several patterns. *J. Appl. Prob.* 22 (1985), 228-234.
222. BREMERMAN, H. L., AND KING, L. L. Determination of codon frequencies and sequence structure of polynucleotides. *Lect. Math. Life Sci.* 5 (1973), 75-110.
223. BRENDL, V. Mapping of transcription terminators of bacteriophages phix174 and g4 by sequence analysis. *J. Virol.* 53 (1985), 340-342.
224. BRENDL, V., BECKMANN, J. S., AND TRIFONOV, E. N. Linguistics of nucleotide sequences: Morphology and comparison of vocabularies. *J. Biomol. Struct. Dyn.* 4 (1986), 11-20.
225. BRENDL, V., AND BUSSE, H. G. Genome structure described by formal languages. *Nucl. Acids Res.* 12 (1984), 2561-2568.
226. BRENDL, V., HAMM, G. H., AND TRIFONOV, E. N. Terminators of transcription with RNA polymerase from *Escherichia coli*: What they look like and how to find them. *J. Biomol. Struct. Dyn.* 3(4) (1986), 705-723.
227. BRENDL, V., AND TRIFONOV, E. N. Computer-aided mapping of DNA-protein sites. *CODATA Bull.* 56 (1984), 17-20.
228. BRENDL, V., AND TRIFONOV, E. N. A computer algorithm for testing potential prokaryotic terminators. *Nucl. Acids Res.* 12(10) (1984), 4411-4427.
229. BRENNER, I., BARRON, S., AND WITTEN, M. A system design approach to a developmental platform for a computational and molecular biology environment. Hawaii Int. Conf. On System Sciences, January 7-10, Kauai, Hawaii, (abstract), 1992.
230. BRESLAUER, D., AND GALIL, Z. An optimal $\mathcal{O}(\log \log n)$ time parallel string matching algorithm. *SIAM J. Computing* 19 (1990), 1051-1058.
231. BRIGGS, F. A. (ed.). *Proceedings of the 1988 International Conference on Parallel Processing*, (The Pennsylvania State University Press, University Park, PA), 1988.
232. BRINKLEY, J. F. Knowledge driven ultrasonic three-dimensional organ modeling. *IEEE Trans. Pattern Analysis and Machine Intelligence PAMI-7(4)* (1985), 431-441.
233. BRINKLEY, J. F. Representing biologic objects as geometric constraint networks. Proc. AAAI Spring Symp. Series: Artificial Intelligence In Medicine, preprint, 1988.
234. BRINKLEY, J. F., ALTMAN, R. B., DUNCAN, B. S., BUCHANAN, B., AND JARDETZKY, O. Heuristic refinement method for the derivation of protein solution structures: Validation on cytochrome b562. *J. Chem. Info. Comput. Sci.* 28(4) (1988), 194-210.
235. BRINKLEY, J. F., PROTHERO, J. S., PROTHERO, J. W., AND ROSE, C. A framework for the design of knowledge-based systems in structural biology. SCAMPC, to appear, 1989.
236. BRINT, A. T., DAVIES, H. M., MITCHELL, E. M., AND WILLETT, P. Rapid geometric searching in protein structures. *J. Mol. Graphics*. 7 (1989), 48-53.
237. BROOKS, L. D., WEIR, B. C., AND SCHAFER, H. E. The probabilities of similarities in DNA sequence comparisons. *Genomics* 3 (1988), 207-216.
238. BROWN, A. P. Evidence for the remnants of an ordered codon sequence and a restricted codon composition in selected proteins. *J. Theor. Biol.* 83 (1980), 537-560.
239. BROWN, C. M., STOCKWELL, P. A., TROTMAN, C. N. A., AND TATE, W. P. Sequence analysis suggests that tetra-nucleotides signal the termination of protein synthesis in eukaryotes. *Nucl. Acids Res.* 18 (1990), 6339-6345.

240. BROWN, J. W. S. A catalogue of splice junction and putative branch point sequences from plant introns. *Nucl. Acids Res.* **14** (1986), 9549-9561.
241. BROWN, N. H. Jr. Neural network implementation approaches for the Connection Machine, (in) **Neural Information Processing Systems**, Anderson, D.Z., (ed.), (American Institute Physics, NY), 1988.
242. BROWNER, M. F., AND LAWRENCE, C. B. Comparative sequence analysis as a tool for studying the secondary structure of mRNAs. *Nucl. Acids Res.* **13(29)** (1985), 8645-8660.
243. BRUCCOLERI, R. E., AND HEINRICH, G. An improved algorithm for nucleic acid secondary structure display. *CABIOS* **4** (1988), 167-173.
244. BRUNAK, S., ENGELBRECHT, J., AND KNUDSEN, S. Cleaning up gene databases. *Nature* **343** (1990), 123.
245. BRUNAK, S., ENGELBRECHT, J., AND KNUDSEN, S. Neural network detects errors in the assignment of mRNA splice sites. *Nucl. Acids Res.* **18** (1990), 4797-4801.
246. BRUNAK, S., ENGELBRECHT, J., AND KNUDSEN, S. Prediction of human mRNA donor and acceptor sites from the DNA sequence. *J. Mol. Biol.* **220** (1991), 49-65.
247. BRUNK, C. F., AND SADLER, L. A. Characterization of the promoter region of tetrahymena genes. *Nucl. Acids Res.* **18** (1990), 323-329.
248. BRUTLAG, D. G., GALPER, A. R., AND MILLIS, D. H. Knowledge-based simulation of DNA metabolism: prediction of enzyme action. *CABIOS* **7(1)** (1991), 9-19.
249. BRUTLAG, D. L. Expert system simulations as active learning environments, (in) **Biomolecular Data: A Resource in Transition**, Colwell, R.R., Swartz, D.G., McDonald, M.T., (eds.), (Oxford University Press, Oxford), pp.185-188, 1989.
250. BRUTLAG, D. L., CLAYTON, J., FRIEDLAND, P., AND KEDES, L. SEQ: A nucleotide sequence analysis and recombination system. *Nucl. Acids Res.* **10** (1982), 279-294.
251. BRUTLAG, D. L., DAUTRICOURT, J. P., MAULIK, S., AND J., R. Sensitive similarity searches of biological sequence databases. *CABIOS* **6(3)** (1990), 237-245.
252. BRUTLAG, D. L., AND KRISTOFFERSON, D. BIONET: an NIH computer resource for molecular biology, (in) **Proceedings of the CODATA Workshop**, May, 1987.
253. BRYANT, S. H. PKB: A program system and data base for analysis of protein structure. *Proteins* **5** (1989), 233-247.
254. BRYANT, S. H., AND STERNBERG, M. J. E. Comparison of protein structural profiles by interactive computer graphics. *J. Mol. Graphics* **5** (1987), 4-7.
255. BUCHER, P. Weight matrix descriptions of four eukaryotic RNA polymerase II promoter elements derived from 502 unrelated promoter sequences. *J. Mol. Biol.* **212** (1990), 563-578.
256. BUCHER, P., AND BRYAN, B. Signal search analysis: A new method to localize and characterize functionally important DNA sequences. *Nucl. Acids Res.* **12** (1984), 287-305.
257. BUCHER, P., AND TRIFONOV, E. N. Compilation and analysis of eukaryotic POL II promoter sequences. *Nucl. Acids Res.* **14(2)** (1986), 10009-10026.
258. BUCHER, P., AND TRIFONOV, E. N. On Nussinov's compilation of eukaryotic transcription initiation sites. *J. Theor. Biol.* **128** (1987), 373-375.
259. BUCHER, P., AND TRIFONOV, E. N. CCAAT box revisited: bidirectionality, location and context. *J. Biomol. Struct. Dyn.* **5** (1988), 1231-1236.
260. BUCKINGHAM, R. H. Codon context. *Experientia* **46** (1990), 1126-1133.
261. BULMER, M. A statistical analysis of nucleotide sequences of introns and exons in human genes. *Mol. Biol. Evol.* **4** (1987), 395-405.
262. BURBAUM, J. J., STARZYK, R. M., AND SCHIMMEL, P. Understanding structure relationships in proteins of unsolved three-dimensional structure. *Proteins: Structure, Function, and Genetics* **7** (1990), 99-111.
263. BURKS, C. The GenBank database and flow of sequences data for the genome, (in) **Biotechnology and The Human Genome: Innovations and Impact**, Woodhead, A.D., Barnhart, B.J., (eds.), (Plenum Publishing Corp., NY), pp.51-56, 1988.
264. BURKS, C. How much sequence data the databanks will be processing in the near future, (in) **Biomolecular Data: A Resource in Transition**, Colwell, R.R., Swartz, D.G., McDonald, M.T., (eds.), (Oxford University Press, Oxford), pp.17-26, 1989.
265. BURKS, C. Sources of data in the GenBank database, (in) **Biomolecular Data: A Resource in Transition**, Colwell, R.R., Swartz, D.G., McDonald, M.T., (eds.), (Oxford University Press, Oxford), pp.327-334, 1989.
266. BURKS, C. The flow of nucleotide sequence data into data banks: Role and impact of large-scale sequencing projects, (in) **Computers and DNA**, SFI Studies in the Sciences of Complexity, Vol. VII, Bell, G., Marr, T., (eds.), (Addison-Wesley, MA), 1990.
267. BURKS, C., FICKETT, J. W., AND GOAD, W. B. GenBank status report. *Science* **235** (1987), 267-268.
268. BURKS, C., FICKETT, J. W., GOAD, W. B., KANEHISA, M., LEWITTER, F. I., RINDONE, W. P., SWINDELL, C. D., TUNG, C.-S., AND BILOFSKY, H. S. The GenBank nucleic acid sequence database. *CABIOS* **1(4)** (1985), 225-233.
269. BURKS, C., LAWTON, J. R., AND BELL, G. I. The LiMB database. *Science* **241** (1988), 888.
270. BURKS, C., AND TOMLINSON, L. Submission of data to GenBank. *Proc. Natl. Acad. Sci.* **86** (1989), 408.

271. BURKS, C. (et. al.). GenBank: Current status and future directions, (in) **Methods in Enzymology-Molecular Evolution: Computer Analysis of Protein and Nucleic Acid Sequences**, Vol.183, Doolittle, R.F., (ed.), (Academic Press, NY) , pp.3-22, 1990.
272. BURNETT, L., BASTEN, A., AND HENSLEY, W. J. An exhaustive tree-searching algorithm for high resolution computer-assisted nucleotide sequence analysis. *CABIOS* 1 (1985), 153-160.
273. BURNETT, L., BASTEN, A., AND HENSLEY, W. J. Should nucleotide sequence analyzing computer algorithms always extend homologies by extending homologies? *Nucl. Acids Res.* 14 (1986), 425-430.
274. BURRIDGE, J. M. An electronic database for molecular biology. *Biochem. Soc. Trans.* 17 (1989), 840-841.
275. BUSETTA, B. Improving the residual representation of proteins. *J. Theor. Biol.* 98 (1982), 621-635.
276. BUSETTA, B. The use of folding patterns in the prediction of protein topologies. *Biochim. Biophys. Acta* 916 (1987), 54-65.
277. BUSETTA, B., AND BARRANS, Y. The prediction of protein topologies. *Biochim. Biophys. Acta* 709 (1982), 73-83.
278. BUSETTA, B., AND BARRANS, Y. The prediction of protein domains. *Biochim. Biophys. Acta* 709 (1984), 117-124.
279. BUSETTA, B., AND HOSPITAL, M. An analysis of the prediction of secondary structures. *Biochim. Biophys. Acta* 701 (1982), 111-118.
280. BUSSARD, A. An international biological data bank: hybridoma data bank. General considerations on these types of banks, (in) **Biotechnology Information '86**, Wakeford, R., (ed.), (IRL Press, Oxford), pp.123-136, 1987.
281. BUTLER, R., BUTLER, T., FOSTER, I., KARONIS, N., OLSON, R., OVERBEEK, R., PFLUGER, N., PRICE, M., AND TUECKE, S. Generating alignments of genetic sequences. *ANL Mathematics and Computer Science Division Technical Memorandum No.132* (1989), ANL/MCS-TM-132.
282. CAHILL, G. F. Jr., AND HINTON, D. R. Howard hughes medical institute and its role in genomic activities. *Genomics* 5 (1989), 952-956.
283. CALLADINE, C. R. Mechanics of sequence-dependent stacking of bases in B-DNA. *J. Mol. Biol.* 161 (1982), 343-352.
284. CAMERON, G., AND KAHN, P. Future nucleotide sequence databases. Macromolecules, Genes and Computers Symposium, (Waterville Valley), NH, August 13-18 (poster only), 1989.
285. CAMERON, G. N. The EMBL data library. *Nucl. Acids Res.* 16(5) (1988), 1865-1867.
286. CAMPIONE-PICCARDO, J. Graphic and DNA sequence interface for the pCP123 database system for managing recombinant DNA data. *CABIOS* 3 (1987), 359-35.
287. CAMPIONE-PICCARDO, J. Algorithms for determining the fate of sites and domain boundaries in computer simulations of recombinant DNA procedures. *CABIOS* 4 (1988), 571-574.
288. CAMPIONE-PICCARDO, J., AND RUBEN, M. An integrated software system for microcomputer management of recombinant DNA data. *Nucl. Acids Res.* 14 (1986), 571-574.
289. CANNON, G. C. Sequence analysis on microcomputers. *Science* 238 (1987), 97-103.
290. CANTOR, C. R. The occurrence of gaps in protein sequences. *biochem. Biophys. Res. Commun.* 31 (1968), 410-416.
291. CAREY, G. F. , (ed). **Parallel Supercomputing: Methods, Algorithms and Applications**, (John Wiley and Sons, NY), 1989.
292. CARHART, R. E., CASH, H. D., AND MOORE, J. F. StrateGene: object-oriented programming in molecular biology. *Nucl. Acids Res.* 12 (1988), 731-737.
293. CARON, P. R. KLONER: A computer program to simulate recombinant DNA strategies by restriction map manipulation. *Nucl. Acids Res.* 12 (1989), 731-737.
294. CARRILLO, H., AND LIPMAN, D. J. The multiple sequence alignment problem in biology. *SIAM J. Appl. Math.* 48 (1988), 1073-1082.
295. CARROLL, M. T., VARGA, G., AND LIM, H. A. Parallel and serial methods of graphically representing polyamino acid sequences with H-trees, FSU-SCRI-91-157, October 1991.
296. CARSON, M., AND BUGG, C. E. Algorithm for ribbon models of proteins. *J. Mol. Graphics* 4 (1986), 121-122.
297. CAVENER, D. R. Comparison of the consensus sequence flanking translational start sites in drosophila and vertebrates. *Nucl. Acids Res.* 15 (1987), 1353-1361.
298. CHAKRAVARTI, A., LASHER, L. K., AND REEFER, J. E. A maximum likelihood method for estimating genome length using genetic linkage data. *Genetics* 128 (1991), 175-182.
299. CHAN, L., AND ZUKER, M. JACOBSON, A. B. A computer method for finding common base paired helices in aligned sequences: application to the analysis of random sequences. *Nucl. Acids Res.* 19(2) (1991), 353-358.
300. CHANG, W. I., AND LAWLER, E. L. Approximate string matching in sublinear expected time, *IEEE Symposium on the Foundations of Computer Science*, 1990.
301. CHEN, G. F. T., AND INOUYE, M. Suppression of the negative effect of minor arginine codons on gene expression and preferential usage of minor codons within the first 25 of the *Escherichia coli* genes. *Nucl. Acids Res.* 18 (1990), 1465-1473.

302. CHEN, H. R., AND BARKER, W. C. The protein identification resource and its applications. *Trends Genetics* **1** (1985), 221-223.
303. CHEN, J.-H., LE, S.-Y., SHAPIRO, B., CURREY, K. M., AND MAIZEL, J. V. A computational procedure for assessing the significance of RNA secondary structure. *CABIOS* **6** (1990), 7-18.
304. CHERFILS, J., BIGRAT, J.-F., LEVIN, J., BATUT, J., AND KAHN, D. Model-building of Fur and FixK DNA-binding domains suggests a basis for specific DNA recognition. *J. Mol. Recognitio*, **2**(3):114-121.
305. CHIARULLI, D. M. A Horizontally Reconfigurable Architecture for Extended Precision Arithmetic, Ph.D thesis, Louisiana State University, 1986.
306. CHIU, D. K. Y., AND KOLODZIEJCZAK, T. Inferring consensus structure from nucleic acid sequences. *CABIOS* **7**(9) (1991), 347-352.
307. CHOTHIA, C. Hydrophobic bonding and accessible surface area in proteins. *Nature* **248** (1974), 338-339.
308. CHOTHIA, C. Structural invariants in protein folding. *Nature* **254** (1975), 304-308.
309. CHOTHIA, C. The nature of the accessible and buried surfaces in proteins. *J. Mol. Biol.* **105** (1976), 1-14.
310. CHOTHIA, C. Principles that determine the structure of proteins. *Ann. Rev. Biochem.* **53** (1984), 537-572.
311. CHOTHIA, C., AND LESK, A. M. Helix movements in proteins. *Trends Biochem. Sci.* **10** (1985), 116-118.
312. CHOTHIA, C., AND LESK, A. M. The relationship between the divergence of sequence and structure in proteins. *EMBO J.* **5** (1986), 823-826.
313. CHOTHIA, C., AND LESK, A. M. The use of sequence homologies to predict protein structures, (in) *Computer Graphics and Molecular Modeling*, Fletterick, R., Zoller, M., (eds.), (Cold Spring Harbor Laboratory, NY), pp.33-37, 1986.
314. CHOU, K.-C., AND CARLACCI, L. Simulated annealing approach to the study of protein structures. *Protein Engineering* **4**(6) (1991), 661-667.
315. CHOU, P. Y., AND FASMAN, G. D. Prediction of protein conformation. *Biochemistry* **13** (1974), 222-245.
316. CHOU, P. Y., AND FASMAN, G. D. Beta-turns in protein. *J. Mol. Biol.* **115** (1977), 135-175.
317. CHOU, P. Y., AND FASMAN, G. D. Prediction of the secondary structure of proteins from their amino acid sequence. *Adv. Enzymol.* **47** (1978), 45-148.
318. CHOU, P. Y., AND FASMAN, G. D. Conservation of chain reversal regions in proteins. *Biophys. J.* **26** (1979), 385-400.
319. CHOU, P. Y., AND FASMAN, G. D. Prediction of beta-turns. *Biophys. J.* **26** (1979), 367-384.
320. CHUPRINA, V. P., LIPANOV, A. A., FEDOROFF, O. Y., AND KIM, S-G. Sequence effects on local DNA topology. *Proc. Natl. Acad. Sci. USA* **88** (1991), 9087-9091.
321. CHURCH, G. M., AND GILBERT, W. Genomic sequencing. *Proc. Natl. Acad. Sci. USA* **81** (1984), 1991-1995.
322. CHURCHILL, G. A. Stochastic models for heterogeneous DNA sequences. *Bull. Math. Biol.* **51**(1) (1989), 79-94.
323. CHURCHILL, G. A., DANIELS, D. L., AND WATERMAN, M. S. The distribution of restriction enzyme sites in *E. coli*. *Nucl. Acids Res.* **18**(3) (1990), 589-597.
324. CHVATAL, V., AND SANKOFF, D. An upper-bound techniques for lengths of common subsequences, (in) *Time Warps, String Edits, and Macromolecules: The Theory and Practice of Sequence Comparison*, Sankoff, D., Kruskal, J.B., (eds.), (Addison-Wesley, Reading), pp.353-357, 1983.
325. CID, H., BUNSTER, M., ARRIAGADA, E., AND CAMPOS, M. Prediction of secondary structure of proteins by means of hydrophobicity profiles. *FEBS Lett.* **150** (1982), 247-254.
326. CIGAN, A. M., AND DONAHUE, T. F. Sequence and structural features associated with translational initiator regions in yeast-a review. *Gene* **59** (1987), 1-18.
327. CINKOSKY, M., NELSON, J. D., AND MARR, T. G. A technical overview of the GenBank/HGIR Database. (manuscript in preparation), 1989.
328. CIPRA, B. Mathematics untwists the double helix. *Science* **247** (1990), 913-915.
329. CLAVERIE, J.-M. A common philosophy and fortran 77 software package for implementing and searching sequence databases. *Nucl. Acids Res.* **12** (1984), 397-407.
330. CLAVERIE, J.-M. The application of sequence data banks, (in) *Biotechnology Information'86*, Wakeford, R., (ed.), (IRL Press, Oxford), pp.49-58, 1987.
331. CLAVERIE, J.-M., AND BOUGUERET, L. Heuristic informational analysis of sequences. *Nucl. Acids Res.* **14** (1986), 179-196.
332. CLAVERIE, J.-M., AND BOUGUERET, L. Neg-entropy and local information content: a new approach to the analysis of nucleotide and amino acid sequences. *Protein, Nucleic Acid and Enzyme* **29** (1986), 111-122.
333. CLAVERIE, J.-M., AND SAUVAGET, I. Assessing the biological significance of primary consensus patterns using sequence databanks. *CABIOS* **1** (1985), 95-104.
334. CLAVERIE, J.-M., SAUVAGET, I., AND BOUGUERET, L. Computer generation and statistical analysis of a data bank of protein sequences translated from GenBank. *Biochimie* **67** (1985), 437-443.
335. CLAVERIE, J.-M., SAUVAGET, I., AND BOUGUERET, L. k-tuple frequency analysis: from intron/exon discrimination to T-cell epitope mapping, (in) *Methods in Enzymology-Molecular Evolution: Computer Analysis of Protein and Nucleic Acid Sequences*, Vol.183, Doolittle, R.F., (ed.), (Academic Press, NY), pp.237-252, 1990.

336. CLAYTON, J., AND KEDES, L. GEL, a DNA sequencing project management system. *Nucl. Acids Res.* 10(1) (1982), 305-321.
337. CLEMENTI, E., CHIN, S., CORONGIU, G., DETRICH, J. H., DUPUIS, M., FOLSOM, D., LIE, G. C., LOGAN, D., AND SONNAD, V. Supercomputing and supercomputers: for science and engineering in general and for chemistry and biosciences in particular. *Biological and Artificial Intelligence Systems*, Clementi, E., Chin, S., (eds.), (ESCOM Science Publishers B.V.: Leiden, The Netherlands), 1988.
338. CLIFT, B., HAUSSLER, D., MCCONNELL, R., SCHNEIDER, T. D., AND STORMO, G. D. Sequence landscape. *Nucl. Acids Res.* 14 (1986), 141-158.
339. COCKWELL, K. Y., AND GILES, I. G. Software tools for motif and pattern scanning: program descriptions including a universal sequence reading algorithm. *CABIOS* 5 (1989), 227-232.
340. COGHLAN, B., AND FRAGA, S. Determination of proteinic structures: an experimentation program. *Computer Physics Communication* 36 (1985), 391-399.
341. COHEN, B. I., PRESNELL, S. R., AND COHEN, F. E. Pattern-based approaches to protein structure prediction, (in) *Methods in Enzymology-Molecular design and Modeling: Concepts and Applications*, Part A, Vol.202, Langone, J.J., (ed.), (Academic Press, NY), pp.252-268, 1991.
342. COHEN, F., WANG, F., LEUNG, M. -Y., HSIU, C. L., AND COHEN, B. Protein structure prediction-methodology and application, (in) *Mathematical Sciences in Genomic Analysis*, in preparation, 1991.
343. COHEN, F. E., ABARBANEL, R. M., KUNTZ, I., AND FLETTERICK, R. Secondary structure assignment for/proteins by a combinatorial approach. *Biochemistry* 22 (1983), 4895-4904.
344. COHEN, F. E., ARBANEL, R. M., KUNTZ, I. D., AND FLETTERICK, R. J. Turn prediction in proteins using a pattern-matching approach. *Biochemistry* 25 (1986), 266-275.
345. COHEN, F. E., AND STERNBERG, M. J. On the prediction of protein structure: The significance of root mean square deviation. *J. Mol. Biol.* 138 (1980), 321-333.
346. COHEN, F. E., STERNBERG, M. J. E., AND TAYLOR, W. R. Analysis and prediction of the packing of alpha-helices against a beta-sheet in the tertiary structure of globular proteins. *J. Mol. Biol.* 156 (1982), 821-862.
347. COLLADO-VIDES, J. A transformational-grammar approach to the study of the regulation of gene expression. *J. Theor. Biol.* 136 (1989), 403-425.
348. COLLADO-VIDES, J. The search for a grammatical theory of gene regulation is formally justified by showing the inadequacy of context-free grammars. *CABIOS* 7(3) (1991), 321-326.
349. COLLINS, J. F., AND COULSON, A. F. W. Applications of parallel processing algorithms for DNA sequence analyses. *Nucl. Acids Res.* 12(1) (1984), 181-192.
350. COLLINS, J. F., AND COULSON, A. F. W. Molecular sequence comparison and alignment, (in) *Nucleic acid and Protein Sequence Analysis: A Practical Approach*, Bishop, M.J., Rawlings, C.J., (eds.), (IRL Press, Oxford), pp.359-385, 1987.
351. COLLINS, J. F., AND COULSON, A. F. W. Significance of protein sequence similarities, (in) *Methods in Enzymology-Molecular Evolution: Computer Analysis of Protein and Nucleic Acid Sequences*, Vol.183, Doolittle, R.F., (ed.), (Academic Press, NY), pp.474-487, 1990.
352. COLLINS, J. F., COULSON, A. F. W., AND LYALL, A. The significance of protein sequence similarities. *CABIOS* 4(1) (1988), 67-71.
353. COLLINS, J. F., AND REDDAWAY, S. F. High-Efficiency Sequence Database Searching: Use of the Distributed Array Processor, (in) *Computers and DNA*, SFI Studies in The Sciences of Complexity, Vol. VII, G. Bell, T. Marr, (eds.), (Addison-Wesley, NY), 1990.
354. COLWELL, R. R. , (ed.). *Biomolecular Data: A Resource in Transition*. (Oxford University Press, NY), 1989.
355. COMAY, E., NUSSINOV, R., AND COMAY, O. An accelerated algorithm for calculating the secondary structure of single stranded RNAs. *Nucl. Acids Res.* 12 (1984), 53-66.
356. CONNELL, C., FUNG, S., HEINER, C., BRIDGHAM, J., CHAKERIAN, V., HERON, E., AND JONES, B. , et al.. Automated DNA sequence analysis. *Biotechniques* 5 (1987), 342-348.
357. CONNOLY, M. L. Analytical molecular surface calculations. *J. Appl. Cryst.* 16 (1983), 548-558.
358. CONNOLY, M. L. Measurements of protein surface shape by solid angles. *J. Mol. Graphics* 4 (1986), 3-6.
359. CONNOLY, M. L. Plotting protein surfaces. *J. Mol. Graphics* 4 (1986), 93-96.
360. CONSEL, C., AND DANVY, O. Partial evaluation of pattern matching in strings. *Info. Proc. Letters* 30 (1990), 79-86.
361. CORBETT, R. J. T., AND ROCHE, R. S. The identification of large peptide fragments produced from proteins of known sequences: a computerized approach using amino acid composition indexes and its application to thermolysin. *Anal. Biochem.* 162 (1987), 546-552.
362. CORE, N. G., EDMISTON, E. W., SALTZ, J. H., AND SMITH, R. M. Parallel Processing of Biological Sequence Comparison Algorithms, 1988. XXX, 266.
363. CORE, N. G., EDMISTON, E. W., SALTZ, J. H., AND SMITH, R. M. Supercomputers and biological sequence comparison algorithms. *Computers Biomed. Res.* 22 (1989), 497-515.
364. CORNETTE, J., AND DELISI, C. Some mathematical aspects of cosmid mapping. *Cell Biophysics* 12 (1988), 271-293.

365. CORNETTE, J. L., CEASE, K. B., MARGALIT, H., SPOUGE, J. L., BERZOFSKY, J., AND DELISI, C. Hydrophobicity scales and computational techniques for detecting amphiphatic structures in proteins. *J. Mol. Biol.* **195** (1987), 659-685.
366. CORNETTE, J. L., AND DELISI, C. The random character of protein evolution and its effect on the reliability of phylogenetic information deduced from amino acid sequences and compositions. *Biochem. J.* **191** (1988), 349-354.
367. CURNISH-BOWDEN, A. Relating proteins by amino acid composition (in) **Methods in Enzymology**, Vol.91, (Academic Press, NY), pp.60-73, 1983.
368. CURNISH-BOWDEN, A. A significance of the purine-pyrimidine motif present in most gene groups. *J. Theor. Biol.* **134** (1988), 1-8.
369. CORPET, F. Multiple sequence alignment with hierarchical clustering. *Nucl. Acids Res.* **16** (1988), 10881-10890.
370. COSIC, I., NESIC, D., PAVLOVIC, D., AND WILLIAMS, R. Enhancer binding proteins predicted by informational spectrum method. *Biochem. Biophys. Res. Commun.* **141** (1986), 831-838.
371. COULSON, A. F. W., COLLINS, J. F., AND LYALL, A. Protein and nucleic acid sequence database searching: a suitable case for parallel processing. *The Computer Journal* **30**(5) (1987), 420-424.
372. COVELLO, P. S., AND GRAY, M. W. Differences in edition at homologous sites in messenger RNAs from angiosperm mitochondria. *Nucl. Acids Res.* **18** (1990), 5189-5196.
373. COWIN, J. E., JELLIS, C. H., AND RICKWOOD, D. A new method of representing DNA sequences which combines ease of visual analysis with machine readability. *Nucl. Acids Res.* **14** (1986), 509-515.
374. CRIPPEN, G. M. Distance geometry and conformational calculations, (in) **Chemometrics Research Studies-1**, Bawden, D., (ed.), (John Wiley, NY), 1981.
375. CROCCHMORE, M. Constant space string matching, (in) **Sequences, Combinatorics, Compression, Security and Transmission**, Capocelli, R.M., (Springer-Verlag), pp.25-40 .
376. CROCCHMORE, M., AND RYTTER, W. Parallel construction of minimal suffix and factor automata. *Info. Proc. Letters* **35** (1990), 121-128.
377. CSANK, C., TAYLOR, F. M., AND MARTINDALE, D. W. Nuclear pre-mRNA introns: analysis and comparison of intron sequences from tetrahymena thermophila and other eukaryotes. *Nucl. Acids Res.* **18** (1990), 5133-5141.
378. CSIRIK, J., AND MAGYAR, J. A computer algorithm to determine the recognition site of restriction enzymes, (in) **Abstracts of: Application of Data Bases and Microcomputers in Molecular Biology**, (Jena), p.36, 1988.
379. CSIRIK, J., MAGYAR, J., AND POLNER, G. A computer algorithm to determine the recognition site of restriction enzymes. *CABIOS* **3**(3) (1987), 245-246.
380. CULL, P., AND HOLLOWAY, J. L. Algorithms for constructing a consensus sequence. Tech. Rep. TR-91-20-1, Oregon State University, Department of Computer Science, 1991.
381. CULLITON, B. J. New game plan for genome mapping. *Science* **245** (1989), 1438-1441.
382. CULLMANN, G., AND LABOUGUES, J. M. The mathematical logic of life. *Origins of Life* **14** (1984), 747-755.
383. CULLMANN, G., AND LABOUGUES, J. M. Evolution of proteins: An ergodic Markov chain. *Math. Mod.* **8** (1987), 639-642.
384. CULLMANN, G., AND LABOUGUES, J. M. The logic of the genetic code. *Math. Mod.* **8** (1987), 643-646.
385. CUOMO, V., MACCHIATO, M. F., AND TRAMONTANO, A. A statistical method for predicting alpha-helical and beta-sheet regions in proteins from their amino acidic sequences. *Il Nuovo Cimento* **3D** (1984), 421-434.
386. CURREY, K., PETERLIN, B. M., AND MAIZEL, J. V. J. Secondary structure of polio virus RNA: Correlation of computer-predicted with electron microscopically observed structure. *Virology* **148** (1986), 33-46.
387. CZELUSNIAK, J., GOODMAN, M., MONCRIEF, N. D., AND KEHOE, S. M. Maximum parsimony approach to construction of evolutionary trees from aligned homologous sequences, (in) **Methods in Enzymology-Molecular Evolution: Computer Analysis of Protein and Nucleic Acid Sequences**, Vol.183, Doolittle, R.F., (ed.), (Academic Press, NY), pp.601-615, 1990.
388. DABROWIAK, J. C., SKOROBAGTY, A., RICH, N., VARY, C. P. H., AND VOURNAKIS, J. Computer assisted microdensitometric analysis of footprinting autoradiographic data. *Nucl. Acids Res.* **14** (1986), 489-499.
389. DANCKAERT, A., CHAPPEY, C., AND HAZOUT, S. 'Size leap' algorithm: an efficient extraction of the longest common motifs from a molecular sequence set. Application to the DNA sequence reconstruction. *CABIOS* **7**(4) (1991), 509-513.
390. DANCKAERT, A., MUGNIER, C., DESSEN, P., AND COHEN-SOLAL, M. A computer program for the design of optimal synthetic oligonucleotides probes for protein coding genes. *CABIOS* **3** (1987), 303-307.
391. DANDEKAR, T., AND SIBBALD, P. R. Trans-splicing of pre-mRNA is predicted to occur in a wide range of organisms including vertebrates. *Nucl. Acids Res.* **18** (1990), 4719-4725.
392. DARDEL, F. Computer simulation of DNA ligation: determination of initial DNA concentrations favouring the formation of recombinant molecules. *Nucl. Acids Res.* **16** (1988), 1767-1778.
393. DARDEL, F., AND BENOUSSAN, P. DNAid: a Macintosh full screen editor featuring a built-in regular expression interpreter for the search of specific patterns in biological sequences using finite state automata. *CABIOS* **4** (1988), 483-486.

394. DARDEN, L. Diagnosing and fixing faults in theories, (in) **Computational Models Of Discovery And Theory Formation**, Shrager, J., Langley, P., (eds.), (Lawrence Erlbaum Associates, Hillsdale, NJ), 1989.
395. DARDEN, T. A method for fitting a smooth ribbon to curved DNA. *J. Computational Chemistry* **10**(4) (1989), 529-551.
396. DAVIES, D. R. A correlation between amino acid composition and protein structure. *J. Mol. Biol* **9** (1964), 437-474.
397. DAVISON, D. Sequence Searching on Supercomputers, (in) **Computers and DNA**, SFI Studies in the Sciences of Complexity, Vol.VII, G. Bell, T. Marr, (eds.), (Addison-Wesley, Reading, MA), 1990.
398. DAVISON, D. B. Sequence similarity ("homology") searching for molecular biologists. *Bull. Math. Biol.* **47** (1985), 437-474.
399. DAVISON, D. B., AND THOMPSON, K. H. A non-metric sequence alignment program. *Bull. Math. Biol.* **46** (1984), 579-590.
400. DAY, G. R., AND BLAKE, R. D. Statistical significance of symmetrical and repetitive segments in DNA. *Nucl. Acids Res.* **10** (1982), 8323-8339.
401. DAY, G. R., AND BLAKE, R. D. Computer analysis and manipulation of DNA sequences. *Computers and Chemistry* **8** (1984), 67-73.
402. DAY, W. H. E. Properties of Levenshtein metrics on sequences. *Bull. Math. Biol.* **46** (1984), 327-332.
403. DAYHOFF, J. E. Distinguished words in data sequences: Analysis and applications to neural coding and other fields. *Bull. Math. Biol.* **46** (1984), 529-543.
404. DAYHOFF, M. O. Computer analysis of protein sequences. *Fed. Proc.* **3** (1974), 2314-2316.
405. DAYHOFF, M. O. Survey of new data and computer methods of analysis, (in) **Atlas of Protein Sequence and Structure** 5, suppl. 3:1-8. (Nat. Biomed. Res. Found., Washington, D.C.), 1978.
406. DAYHOFF, M. O., BARKER, W. C., AND HUNT, L. T. Establishing homologies in protein sequences, (in) **Methods in Enzymology**, Vol.91, (Academic Press, NY), pp.524-544, 1983.
407. DAYHOFF, M. O., AND ECK, R. V. Computer aids to protein sequence determination. *J. Theor. Biol.* **8** (1964), 97-112.
408. DAYHOFF, M. O., AND ECK, R. V. MASSPEC: A computer program for complete sequence analysis of large proteins by mass spectrophotometry data of a single sample. *Comp. Biol. Med.* **1** (1970), 5-28.
409. DAYHOFF, M. O., ECK, R. V., AND PARK, C. M. A model of evolutionary change in proteins, (in) **Atlas of Protein Sequence and Structure**, 5:89-99. (Nat. Biomed. Res. Found., Washington D.C.), 1972.
410. DAYHOFF, M. O., AND ORCUTT, B. C. Methods for identifying proteins by using partial sequences. *Proc. Natl. Acad. Sci. USA* **76** (1979), 2170-2174.
411. DAYHOFF, M. O., SCHWARTZ, R. M., CHEN, H. R., BARKER, W. C., HUNT, L. T., AND ORCUTT, B. C. Nucleic acid sequence database. *DNA* **1** (1981), 51-58.
412. DAYHOFF, M. O., SCHWARTZ, R. M., AND ORCUTT, B. C. A model of evolutionary change in proteins, (in) **Atlas of Protein Sequence and Structure**, 5, suppl. 3:345-352. (Nat. Biomed. Res. Found., Washington D.C.), 1978.
413. DAYRINGER, H., TRAMONTANO, A., AND FLETTERICK, R. J. PROTEUS: graphics software for proteins, (in) **Computer Graphics and Molecular Modeling**, Fletterick, R., Zoller, M., (eds.), (Cold Spring Harbor Laboratory, NY), pp. 5-8, 1986.
414. DEININGER, P., AND SMITH, W. Voice input of DNA sequence data. *Biotechniques* **5** (1987), 6-7.
415. DELBRUCK, M. Knotting problems in biology. *Proc. Symp. Appl. Math.* **14** (1962), 55-63.
416. DELCOIGNE, A., AND HANSEN, P. Sequence comparison by dynamic programming. *Biometrika* **62** (1975), 661-664.
417. DELEAGE, G., AND ROUX, B. An algorithm for protein secondary structure prediction based on class prediction. *Protein Engineering* **1** (1987), 289-294.
418. DELISI, C. Assessing the significance of local sequence homologies. *Math. Biosciences* **69** (1984), 77-85.
419. DELISI, C. An overview of the human genome project, Barnhart, B. (ed.), (Academic Press, NY), 1988.
420. DELISI, C. Computers in molecular biology: Current applications and emerging trends. *Science* **240** (1988), 47-52.
421. DELISI, C. The human genome project. *The American Scientist* **76** (1988), 488-493.
422. DELISI, C. Computation and the human genome project, (in) **Computers and DNA**, Bell, G. and Marr, T. (eds.), (Addison Wesley, NY), pp.13-19, 1990.
423. DELISI, C., AND BERZOFSKY, J. A. T-cell antigenic sites tend to be amphipathic structures. *Proc. Natl. Acad. Sci. USA* **82** (1985), 7048-7052.
424. DELISI, C., KLEIN, P., AND KANEHISA, M. Some comments on protein taxonomy: procedures for functional and structural classification, (in) **Molecular Basis of Cancer, Part A: Macromolecular Structure, Carcinogens, and Oncogenes**, (Alan R. Liss), 1985.
425. DELISI, C., AND VAJDA, S. Structural biology and the human genome project, (in) **Structure and Methods**, Sarma, R.H., and Sarma, M.H. (eds.), (Adenine Press, Schenectady, NY), pp.7-14, 1990.
426. DELISI, D. Statistical thermodynamics of oligomer-polymer interactions. *Biopolymers* **13** (1974), 2305.
427. DELORME, M.-O., AND HENAUT, A. Merging of distance matrices and classification by dynamic clustering. *CABIOS* **4** (1988), 453-458.

428. DENNINGHOFF, K. L., AND GATTERDAM, R. W. On the undecidability of splicing systems. *Int. J. Computer Math.* **27** (1989), 133-145.
429. DEVEREUX, J. A rapid method for identifying sequences in large nucleotide sequence databases. Ph.D. Thesis, University of Wisconsin, 1988.
430. de Vlieg, J., AND van Gunsteren, W. F. Combined procedures of distance geometry and molecular dynamics for determining protein structure from nuclear magnetic resonance data, (in) *Methods in Enzymology-Molecular design and Modeling: Concepts and Applications*, Part A, Vol.202, Langone, J.J., (ed.), (Academic Press, NY), pp.268-300, 1991.
431. DEWACHTER, R. The number of repeats expected in random nucleic acid sequences and found in genes. *J. Theor. Biol.* **91** (1981), 71-98.
432. DIAL, R., GLOVER, F., KARNEY, D., AND KLINGMAN, D. A computation analanalysis of alternative algorithms and labelling techniques for finding shortest path trees. *Networks* **9** (1979), 215-248.
433. DIAMOND, R. Applications of computer graphics in molecular biology. *Computer Graphics Forum* **3** (1984), 3-11.
434. DICKERSON, R. E., DREW, H. R., CONNER, B. N., KOPKA, M. L., AND PJURA, P. E. Helix geometry and hydration in A-DNA, B-DNA, and Z-DNA. Cold SpringHarbor Symp. *Computer Graphics Forum* **47** (1983), 13-24.
435. DICKERSON, R. E., DREW, H. R., CONNOR, B. N., WING, R. M., FRATINI, A. V., AND KOPKA, M. L. The anatomy of A-, B- and Z-DNA. *Science* **216** (1982), 475-485.
436. DICKERSON, R. E., KOPKA, M. L., AND PJURA, P. E. A random-walk model for helix bending in B-DNA. *Proc. Natl. Acad. Sci. USA* **80** (1983), 7099-7103.
437. DICKMAN, S. New protein database for europe. *Nature* **327** (1987), 265.
438. DIX, T. I., AND KIERONSKA, D. H. Errors between sites in restriction site mapping. *CABIOS* **4** (1988), 117-124.
439. DODD, I. B., AND EGAN, J. B. Improved detection of helix-turn-helix DNA-binding motifs in protein sequences. *J. Mol. Biol.* Submitted.
440. DODD, I. B., AND EGAN, J. B. Systematic method for the detection of potential lambda cro-like DNA-binding regions in proteins. *J. Mol. Biol.* **194** (1987), 557-564.
441. DOE Human Genome Steering Committee Report, October 18, 1988.
442. DOE Human Genome Steering Committee Report, January 16, 1989.
443. DOE/NIH Human Genome Contractors/Grantee Workshop, 1989.
444. DONGARRA, J. J., AND DUFF, I. S. Advanced Architecture Computers, ANL/MCS-TM-57, Rev. 2, Argonne National Laboratory, 1989.
445. DOOLITTLE, R. F. Protein evolution, (in) *The Proteins*, Neurath, H., Hills, R.L., (eds.), 2nd ed., 4:1-118, (Academic Press, NY), 1979.
446. DOOLITTLE, R. F. Similar amino acid sequences: chance or common ancestry. *Science* **214** (1981), 149-159.
447. DOOLITTLE, R. F. Of URFs and ORFs: a primer on how to analyze derived amino acid sequences, (in) *University Science Books*, (Mill Valley, California), 1986.
448. DOOLITTLE, R. F. Searching through sequence databases, (in) *Methods in Enzymology-Molecular Evolution: Computer Analysis of Protein and Nucleic Acid Sequences*, Vol.183, Doolittle, R.F., (ed.), (Academic Press, NY), pp.99-110, 1990.
449. DOOLITTLE, R. F., FENG, D. F., AND JOHNSON, M. S. Computer-based characterization of epidermal growth factor precursor. *Nature* **307** (1984), 558-560.
450. DOOLITTLE, R. F., AND FENG, D.-F. Nearest neighbor procedure for relating progressively aligned amino acid sequences, (in) *Methods in Enzymology-Molecular Evolution: Computer Analysis of Protein and Nucleic Acid Sequences*, Vol.183, Doolittle, R.F., (ed.), (Academic Press, NY), pp.659-669, 1990.
451. DOPAZO, J. Reducing the effect of the data order in algorithms for constructing phylogenetic trees. *CABIOS* **4** (1988), 307.
452. DOUBET, S., BOCK, K., SMITH, D., DARVILL, A., AND ALBERSHEIM, P. The complex carbohydrate structure database. *Trends Biochem. Sci.* **14** (1989), 475-477.
453. DOUTHART, R. J., SCHMALTZ, J. E., AND THOMAS, J. J. Color graphic representations of large sequences in the GEM environment. *Nucl. Acids Res.* **16** (1988), 1657-1666.
454. DOUTHART, R. J., THOMAS, J. J., ROSIER, S. D., SCHMALTZ, J. E., AND WEST, J. W. Cloning simulation in the CAGE environment. *Nucl. Acids Res.* **14** (1986), 285-297.
455. DRESS, A., VON HAESELER, A., AND KRUEGER, M. Reconstructing phylogenetic trees using variants of the "four-point-condition". *Studien zur Klassifikation* **17** (1986), 299-305.
456. DUBINER, M., GALIL, Z., AND MAGEN, E. Faster tree pattern matching. May be a technical report.
457. DUBNICK, M., AND MOUNT, D. W. SPLICE, a computer program for automated extraction of information from Genbank sequence entries. *CABIOS* **6(4)** (1990), 343-346.
458. DUGGLEBY, R. G. A paucity of palindromes in phix174. *J. Theor. Biol.* **93** (1981), 143-155.
459. DUMAS, J. P., AND NINIO, J. Efficient algorithms for folding and comparing nucleic acid sequences. *Nucl. Acids Res.* **10(1)** (1982), 197-206.

460. DUNCAN, B. S. Computation of Protein Structures from Experimental Data, Ph.D. Thesis, Stanford University Biophysics Department, 1989.
461. DUNN, S. M., FEDYNA, J., AND PETERS, J. E. Efficient parallel implementation of sequence analysis algorithms using a global address space model. Advances in Mathematics in Computers and Medicine: Mathematical Models in Medicine, 1990. (in press).
462. DUNNILL, P. The use of helical net-diagrams to represent protein structures. *Biophys. J.* 8 (1968), 865-875.
463. DURAND, R., AND BREGEGERE, F. An efficient program to construct restriction maps from experimental data with realistic error levels. *Nucl. Acids Res.* 12 (1984), 703-716.
464. EASTHOPE, P. L., AND HAVEL, T. F. Computational experience with an algorithm for the tetrangle inequality bound smoothing. *Bull. Math. Biol.* 51(1) (1989), 173-194.
465. EBELING, W., AND JIMENEZ-MONTANO, M. A. On grammars, complexity, and information measures of biological macromolecules. *Math. Bio. Sci.* 52 (1980), 53-71.
466. ECCLES, J. R., AND SALDANHA, J. W. Megadata-based generation and management of knowledge-bases from molecular biological databases. *Computer Methods and Programs in Biomedicine* 32 (1990), 115-124.
467. ECK, R. V. A simplified strategy for sequence analysis of large proteins. *Nature, Lond.* 193 (1962), 241-243.
468. EDELSBRUNNER, H. A new approach to rectangle intersections, Parts 1 and 2. *Int. J. Comp. Math.* 13 (1983), 209-229.
469. EDWARDS, M. S., STERNBERG, M. J. E., AND THORNTON, J. M. Structural and sequence patterns in loops of beta-alpha-beta units. *Protein Engineering* 1 (1987), 173-181.
470. EHRLICH, R., MARIN, M., LAROUSSE, A., GABARRO-ARPA, J., SCHMITT, B., AND REISS, C. Promoter recognition and transcription initiation in *E. coli*. *Folia Biologica* 30 (1984), 105-118.
471. EIGEN, M., LINDEMANN, B. F., TIETZE, M., WINKLER-OZWATITSCH, R., DRESS, A., AND VON HAESELER, A. How Old is the Genetic Code? Statistical Geometry of tRNA Provides an Answer. *Science* 244 (1989), 673-679.
472. EIGEN, M., AND WINKLER-OZWATITSCH, R. Statistical geometry on sequence space, (in) *Methods in Enzymology, Molecular Evolution: Computer Analysis of Protein and Nucleic Acid Sequences*, Vol.183, Doolittle, R.F., (ed.), (Academic Press, NY), pp.505-530, 1990.
473. EILAM-TZOREFF, T., AND VISHKIN, U. Matching patterns in strings subject to multi-linear transformations. *Theoretical Computer Science* 60 (1988), 231-254.
474. EISENBERG, D. Three-dimensional structure of membrane and surface proteins. *Ann. Rev. Biochem* 53 (1986), 595-623.
475. EISENBERG, D., AND McLACHLAN, A. D. Solvation energy in protein folding and binding. *Nature* 319 (1986), 199-203.
476. EISENBERG, D., SCHWARZ, E., KOMAROMY, M., AND WALL, R. Analysis of membrane and surface protein sequences with the hydrophobic moment plot. *J. Mol. Biol.* 179 (1984), 125-142.
477. EISENBERG, D., WEISS, R. M., AND TERWILLIGER, T. C. The helical hydrophobic moment: a measure of the amphilicity of a helix. *Nature* 299 (1982), 371-374.
478. EISENBERG, D., WEISS, R. M., AND TERWILLIGER, T. C. The hydrophobic moment detects periodicity in protein hydrophobicity. *Nat. Acad. Sci. USA* 81 (1984), 140-144.
479. EISENBERG, D., WEISS, R. M., TERWILLIGER, T. C., AND WILCOX, W. Hydrophobic moments and protein structure. *Faraday Symposia Chemical Society* 17 (1982), 109-120.
480. EISENBERG, D., WILCOX, W., AND ESHITA, S. Free energy and protein folding: hydrophobic moments and solvation free energy, (in) *Computer Graphics and Molecular Modeling*, Fletterick, R., Zoller, M., (eds.), (Cold Spring Harbor Laboratory, NY), pp.59-65, 1986.
481. EISENBERG, D., WILCOX, W., AND ESHITA, S. Hydrophobic moments as tools for analysis of protein sequences, (in) *Proteins: Structure and Function*, L'Italien J.J., (ed.), (Plenum, NY), pp.425-436, 1987.
482. ELDER, J. K., GREEN, D. K., AND SOUTHERN, E. M. Automatic reading of DNA sequencing gel autoradiographs using a large format digital scanner. *Nucl. Acids Res.* 14 (1986), 417-424.
483. ELDER, J. K., AND SOUTHERN, E. M. Automated reading of DNA sequencing gels autoradiographs, (in) *Nucleic Acid and Protein Sequence Analysis: A Practical Approach*, Bishop, M.J., Rawlings, C.J., (eds.), (IRL Press, Washington, D.C.), 1987.
484. ELLEMAN, T. C. A method for detecting distant evolutionary relationships between protein or nucleic acid sequences in the presence of deletions or insertions. *J. Mol. Evol.* 11 (1978), 143-161.
485. ELTON, R. Theoretical models for heterogeneity of base composition in DNA. *J. Theor. Biol.* 45 (1974), 533-553.
486. ENGEL, L. W. Database management for a recombinant DNA bank. *Bio/Technology* 3 (1985), 329-335.
487. ENGELMAN, D. M., STEITZ, T. A., AND GOLDMAN, A. Identifying nonpolar transbilayer helices in amino acid sequences of membrane proteins. *Ann. Rev. Biophys. Biophys. Chem.* 15 (1986), 321-353.
488. ENGEN, S., AND SAETHER, B.-E. Optimization of constrained life histories: some generalizations. *J. Theor. Biol.* 130 (1988), 229-237. (abstract only).

489. EPPIG, J. T., REINER, A. H., MOBRAATEN, L. E., SNELL, T. C., AND NADEAU, J. H. Multi-level analysis and display of mouse genome data, *Macromolecules, Genes, and Computers Symposium*, (Waterville Valley, NH), August pp.13-18, (abstract only), 1989.
490. EPPSTEIN, D. Sequence comparison with mixed convex and concave costs. *J. Algorithms* 11(1) (1990), 85-1101.
491. EPPSTEIN, D., AND GALIL, Z. Parallel algorithmic techniques for combinatorial computation. *Annual Review of Computer Science* 3 (1988), 233-283.
492. EPPSTEIN, D., GALIL, Z., AND GIANCARLO, R. Speeding up Dynamic Programming, *IEEE Symposium on the Foundations of Computer Science* (1988) 488-496.
493. EPPSTEIN, D., GALIL, Z., AND GIANCARLO, R. Speeding up dynamic programming with application to the computation of RNA structure. preprint.
494. EPPSTEIN, D., GALIL, Z., AND GIANCARLO, R. Efficient algorithms with applications to molecular biology, (in) *Sequences, Combinatorics, Compression, Security and Transmission*, Capocelli, R.M. (Springer-Verlag), pp.59-74, 1990.
495. EPPSTEIN, D., GALIL, Z., GIANCARLO, R., AND ITALIANO, G. F. Sparse dynamic programming I (concave/convex cost functions), *J. ACM*, in press.
496. EPPSTEIN, D., GALIL, Z., GIANCARLO, R., AND ITALIANO, G. F. Sparse dynamic programming I (linear cost functions), *J. ACM*, in press.
497. EPPSTEIN, D., GIANCARLO, R., GALIL, Z., AND ITALIANO, F. Sparse dynamic programming. preprint.
498. ERDMANN, V. A., AND WOLTERS, J. The Berlin RNA databank. *Protein Seq. Data Anal.* 1 (1987), 127.
499. ERDMANN, V. A., WOLTERS, J., DIGWEED, M., PIELER, T., LINDSCHAU, C., S., L., AND ULRICH, N. Computer analysis of ribosomal 5S RNA sequences and their correlation with biochemical structural data, (in) *Computer Handling and Dissemination of Data*, Glaeser, P.S., (ed.), (Elsevier, North Holland), pp.373-380, 1987.
500. ERICKSON, B. W., AND ALTSCHUL, S. F. Significance of DNA or RNA alignments: random permutation methods. *Fed. Proc.* 44 (1985), 1086.
501. ERICKSON, B. W., SEKULSKI, J. M., AND SELLERS, P. H. Recognition of patterns in protein sequences by metric analysis, (in) *Methods in Protein Sequence Analysis*, Elzinga, M., (ed.), (Humana Press, Clifton, NJ), pp.553-554, 1982.
502. ERICKSON, B. W., AND SELLERS, P. H. Recognition of patterns in genetic sequences, (in) *Time Warps, String Edits, and Macromolecules: The Theory and Practice of Sequence Comparison*, David Sankoff, Joseph B. Kruskal, (eds.), (Addison-Wesley, Reading, Massachusetts), 1983.
503. ERICKSON, J. W., AND ALTMAN, G. G. A search for patterns in the nucleotide sequence of the MS2 genome. *J. Math. Biol.* 7 (1979), 219-230.
504. ERIKSON, B. W., SEKULSKI, J. M., AND SELLERS, P. H. Recognition of patterns in protein sequences, (in) *Methods in Protein SequenceAnalysis*, Elzinga, M., (ed.), (Humana Press, Clifton, NJ), 1982.
505. ERNST, C., AND SUMNERS, D. W. A calculus for rational tangles: applications for DNA recombination, preprint, 1990.
506. EZHOV, A. A., KALAMBET, Y. A., AND CHERNY, D. I. Neuron network for the recognition of *E. coli* promoters. *Studia Biophysica* 129 (1989), 183-192.
507. EZHOV, A. A., KALAMBET, Y. A., AND CHERNY, D. Y. Neuron net for the recognition of *E. coli* promoters, (in) *Abstracts of: Application of Data Bases and Microcomputers in Molecular Biology*, (Jena), p.30, 1990.
508. FACCHIANO, A., FACCHIANO, A. M., FACCHIANO, F., RAGONE, R., AND COLONNA, G. FAST (Flexible Analysis by Software Tool) and CHAMP (Chemico-physical amino acidic parameter data bank): a new tool to investigate protein structure. *CABIOS* 5 (1989), 299-303.
509. FACCHIANO, F., FACCHIANO, A., FACCHIANO, A. M., RAGONE, R., AND COLONNA, G. New graphic representation of structural parameters of proteins. *CABIOS* 4 (1988), 303-305.
510. FANDIC, M. Fragment search in acyclic structures. *J. Chem. Inf. Comput. Sci.* 18 (2) (1978), 101-107.
511. FARBER, R., LAPEDES, A., AND SIROTKIN, K. Determination of eukaryotic protein coding regions using neural networks and information theory. Tech. Rep. LA-UR-90-4014, Los Alamos National Laboratory, Theoretical Division, MS B213, Los Alamos National Laboratory, Los Alamos, NM, 87545, 1991.
512. FARRIS, J. J. Estimation of the number of amino acid substitutions when back mutations can occur. *Am. Naturalist* 107 (1973), 531-534.
513. FASMAN, G. D. The prediction of the secondary structure of proteins, (in) *Methods in Protein Sequence Analysis*, Jornvall, H., Hoog, J.-O., and Gustavsson, A.-M., (ed.), (Birkhauser Verlag, Boston, MA), pp.321-332, 1991.
514. FASMAN, G. D., AND GILBERT, W. A. The prediction of transmembrane protein sequences and their conformation: an evaluation. *Trends Biochem. Sci.* 15 (1990), 89-92.
515. FAUCHERE, J. L., AND PLISKA, V. E. Hydrophobic parameters pi of amino acid side chains from the partitioning of N-acetyl-amino acid amides. *Eur. J. Med. Chem.* 18 (1983), 369-375.
516. FAULKNER, D. V., AND JURKA, J. Multiple Aligned Sequence Editor, Molecular Biology Computer Research Resource Document, Release 2.3, 1989.

517. FELDMANN, R. J. The design of computing systems for molecular modeling. *Ann. Rev. Biophys. Bioeng.* 5 (1976), 477-510.
518. FELDMANN, R. J., BING, D. H., POTTER, M., MAINHART, C., FURIE, B., FURIE, C., AND CAPORALE, L. H. On the construction of computer models of proteins by the extension of crystallographic structures. *Ann. NY. Acad. Sci.* 439 (1985), 12-43.
519. FELSENSTEIN, J. Evolutionary trees from DNA sequences: a maximum likelihood approach. *J. Mol. Evol.* 17 (1981), 368-376.
520. FELSENSTEIN, J. Numerical methods for inferring evolutionary trees. *Quat. Rev. Biol.* 57 (1982), 379-404.
521. FELSENSTEIN, J. Phylogenies from gene frequencies: a statistical problem. *Syst. Zool.* 34 (1985), 300-311.
522. FELSENSTEIN, J. Phylogenies from molecular sequences: inference and reliability. *Ann. Rev. Genet.* 22 (1988), 521-565.
523. FELSENSTEIN, J. S., SAWYER, S., AND KOCHIN, R. An efficient method for matching nucleic acid sequences. *Nucl. Acids Res.* 10(1) (1982), 133-139.
524. FENG, D. F., AND DOOLITTLE, R. F. Progressive sequence alignment as a prerequisite to correct phylogenetic trees. *J. Mol. Evol.* 25 (1987), 351-360.
525. FENG, D. F., JOHNSON, M. S., AND DOOLITTLE, R. F. Aligning amino acids sequences: comparison of commonly used methods. *J. Mol. Evol.* 21 (1985), 112-125.
526. FENG, D-F., AND DOOLITTLE, R. F. Progressive alignment and phylogenetic tree construction of protein sequences, (in) *Methods in Enzymology-Molecular Evolution: Computer Analysis of Protein and Nucleic Acid Sequences*, Vol.183, Doolittle, R.F., (ed.), (Academic Press, NY), pp.375-387, 1990.
527. FERGUSON-SMITH, A. C., FIENBERG, A., AND RUDDLE, R. H. Isolation, chromosomal localization, and nucleotide sequence of the human HOX 1.4 Homeobox. *Genomics* 5 (1989), 250-258.
528. FERRAN, E. A., AND FERRARA, P. Topological maps of protein sequences. *Biol. Cybern.* 65 (1991), 451-458.
529. FERRAN, E. A., AND FERRARA, P. Clustering proteins into families using artificial neural networks. *CABIOS* 8(1) (1992), 39-44.
530. FERRETTI, V., AND SANKOFF, D. A continuous analog for RNA folding. *Bull. Math. Biol.* 51(1) (1989), 167-171.
531. FERRIN, T. E., HUANG, C. C., JARVIS, L. E., AND LANGRIDGE, R. The MIDAS database system. *J. Mol. Graphics* 6 (1988), 2-12.
532. FERRIN, T. E., HUANG, C. C., JARVIS, L. E., AND LANGRIDGE, R. The MIDAS display system. *J. Mol. Graphics* 6 (1988), 13-27.
533. FERTIG, S., AND GELERNTER, D. H. A virtual machine for database-driven expert systems, (in) *Proc. IEEE Conf. AI Tools and Architectures*, (Fairfax, Virginia), 1989. in press.
534. FICHANT, G., AND GAUTIER, C. Statistical method for predicting protein coding regions in nucleic acid sequences. *CABIOS* 3(3) (1987), 287-295.
535. FICKETT, J. Fast optimal alignment. *Nucl. Acids Res.* 12(1) (1984), 175-179.
536. FICKETT, J. W. Recognition of protein coding regions in DNA sequences. *Nucl. Acids Res.* 10 (1982), 5303-5318.
537. FICKETT, J. W. Correct transmission of protein coding regions in GenBank. *Trends Biochem. Sci.* 11 (1986), 190.
538. FICKETT, J. W., AND BURKS, C. Development of a database for nucleotide sequences, (in) *Mathematical Methods For DNA Sequences*, M. Waterman, (ed.), (CRC Press, Boca Raton), 1989.
539. FIELDING, A. H., AND SMITH, C. A. A rapid computer technique for analysing molecular interactionsa rapid computer technique for analysing molecular interactions. *CABIOS* 2(3) (1986), 195-200.
540. FIELDS, C. Information content of *caenorhabditis elegans* splice site sequences varies with intron length. *Nucl. Acids Res.* 18 (1990), 1509-1512.
541. FIGUREAU, A. Optimization and the genetic code. *Origins of Life* 19 (1989), 57-68.
542. FIGUREAU, A., AND LABOUGUES, J. M. Quantitative models for the evolution of the genetic code, (in) *Proc. 1st Int. Conf. Applied Modelling and Simulation*, (Lyon, France), Sept.7-11, Vol.V, pp.10-14, 1981.
543. FIJKES, R., AND KEHLER, T. The role of frame-based representation in reasoning. *CACM* 28, 9 (1985), 904-920.
544. FILIPSKI, J. Correlation between molecular clock ticking, codon usage, fidelity of DNA repair, chromosome banding and chromatin compactness in germline cells. *FEBS Lett.* 217 (1987), 184-186.
545. FILIPSKI, J. Why the rate of silent codon substitutions is variable within a vertebrate's genome. *J. Theor. Biol.* 134 (1988), 159-164.
546. FILIPSKI, J., SALINAS, J., AND RODIER, F. Two distinct compositional classes of vertebrate gene-bearing DNA stretches, their structures and possible evolutionary origin. *DNA* 6 (1987), 109-118.
547. FILIPSKI, J., (et. al.). Periodicity of DNA folding in higher order chromatin structures. *EMBO J.* 9 (1990), 1319-1328.
548. FISCHEL-GHODSIAN, F., MATHIOWITZ, G., AND SMITH, T. F. Alignment of protein sequences using secondary structure: a modified dynamic programming method. *Protein Engineering* 3(7) (1990), 577-581.

549. FISHLEIGH, R. V., ROBSON, B., GARNIER, J., AND FINN, P. W. Studies on rationales for an expert system approach to the interpretation of protein sequence data. *FEBS Lett.* **214** (1987), 219–225.
550. FITCH, W. M. An improved method for testing for evolutionary homology. *J. Mol. Biol.* **16** (1966), 9–16.
551. FITCH, W. M. Locating gaps in amino acid sequences to optimize the homology between two proteins. *Biochem. Genet.* **3** (1969), 99–108.
552. FITCH, W. M. Further improvements in the method of testing for evolutionary homology among proteins. *J. Mol. Biol.* **49** (1970), 1–14.
553. FITCH, W. M. Toward defining the course of evolution: minimum change for a specific tree topology. *Syst. Zool.* **20** (1971), 406–416.
554. FITCH, W. M. On the problem of discovering the most parsimonious tree. *Am. Naturalist* **111** (1975), 223–257.
555. FITCH, W. M. A non-sequential methods for constructing trees and hierarchical classifications. *J. Mol. Evol.* **18** (1981), 30.
556. FITCH, W. M. The old REH theory remains unsatisfactory and the new REH theory is problematical—a reply to Holmquist and Jukes. *J. Mol. Evol.* **18** (1981), 60–67.
557. FITCH, W. M. Calculating the expected frequencies of potential secondary structure in nucleic acids as a function of stem length, loop size, base composition and nearest-neighbor frequencies. *Nucl. Acids Res.* **11** (1983), 4655–4663.
558. FITCH, W. M. Random sequences. *J. Mol. Biol.* **163** (1983), 171–176.
559. FITCH, W. M., AND LANGLEY, C. H. Protein evolution and molecular clock. *Fed. Proc.* **35** (1976), 2092–2097.
560. FITCH, W. M., AND MARGOLIASH, W. Construction of phylogenetic trees. *Science* **155** (1967), 279.
561. FITCH, W. M., AND MARKOWITZ, E. An improved method for determining codon variability in a gene and its application to the rate of fixation of mutations in evolution. *Biochem. Genet.* **4** (1970), 579–593.
562. FITCH, W. M., AND SMITH, T. F. Implications of minimal length trees. *Syst. Zool.* **31** (1982), 68–75.
563. FITCH, W. M., AND SMITH, T. F. Optimal sequence alignments. *Proc. Natl. Acad. Sci. USA* **80** (1983), 1382–1386.
564. FITCH, W. M., SMITH, T. F., AND RALPH, W. W. Mapping the order of DNA restriction fragments. *Gene* **22** (1983), 19–29.
565. FLOEGEL, R., ZIELENKIEWICZ, AND SAENGER, W. Recognition pattern of different bases in the active site of ribonuclease Ms-A model building study. *J. Biomole. Struct. Dyn.* **7**(2) (1989), 257–268.
566. FOLEY, B. T., NELSON, D., SMITH, M. T., AND BURKS, C. Cross-sections of the GenBank database. *Trends in Genetics* **2**(9) (1986), 233–238.
567. FOLZ, R. J., AND GORDON, J. I. Computer-assisted predictions of signal peptidase processing sites. *Biochem. Biophys. Res. Commun.* **146** (1987), 870–877.
568. FOLZ, R. J., AND GORDON, J. I. PARA-SITE: a computer algorithm for rapidly analyzing the physical-chemical properties of amino acid sequences at sites of co-and post-translational protein processing. *CABIOS* **4** (1988), 175–179.
569. FONDRET, C., DESSEN, P., AND LE BEUX, P. Principle of codification for quick comparisons with the entire biomolecule databanks and associated programs in FORTRAN 77. *Nucl. Acids Res.* **14** (1986), 197–204.
570. FORBUS, K. Qualitative process theory. Tech. Rep. TR-789, MIT AI Laboratory, 1984.
571. FOULSER, D. E. A linear time algorithm for DNA sequencing. Tech. Rep. YALEU/DCS/RR-812, Yale University, Department of Computer Science, 1990.
572. FOULSER, D. E., AND CORE, N. G. Parallel computation of multiple biological sequence comparisons. *Computers and Biomedical Research* **23** (1990), 310–331.
573. FRAGA, S. Theoretical prediction of protein antigenic determinants from amino acid sequences. *Can. J. Chem.* **60** (1982), 2606–2610.
574. FRAGA, S., SINGH, B., AND COGHLAN, B. Theoretical studies of protein structures: prediction of antigenic determinants. *J. Mol. Structure* **120** (1985), 213–226.
575. FREDMAN, M. L. Algorithms for computing evolutionary similarity measures with length dependent gap penalties. *Bull. Math. Biol.* **46** (1984), 553–566.
576. FREIER, S. M., KIERZEK, R., JAEGER, J. A., SUGIMOTO, N., CARUTHERS, M. H., NEILSON, T., AND TURNER, D. H. Improved free-energy parameters for predictions of RNA duplex stability. *Proc. Natl. Acad. Sci. USA* **83** (1986), 9373–9377.
577. FREIER, S. M., SINCLAIR, A., NEILSON, T., AND TURNER, D. H. Improved free energies for G.C base-pairs. *J. Mol. Biol.* **185** (1985), 645–647.
578. FRENKEL, K. The human genome project and informatics. *CACM* **34**(11) (1991), 41–51.
579. FRIEDEMANN, T. Alignment of multiple DNA and protein sequence display. *CABIOS* **4** (1988), 213–214.
580. FRIEDLAND, P. Knowledge-based Hierarchical Planning in Molecular Genetics, Ph.D. Thesis, Stanford University Computer Science Department, Report CS-79-760, September, 1979.
581. FRIEDLAND, P., AND KEDES, L. H. Discovering the secrets of DNA. *CACM* **28** (11) (1985), 1164–1188.

582. FRIEDLAND, P. E., AND IWASKI, Y. The concept and implementation of skeletal plans. *J. Automated Reasoning* 1 (1985), 161-208.
583. FRIEDRICH, M. S., AND WOLYNES, P. G. Toward protein tertiary structure recognition by means of associative memory hamiltonians. *Science* 246 (1989), 371-373.
584. FRISTENSKY, B. Improving the efficiency of dot-matrix similarity searches through use of oligomer table. *Nucl. Acids Res.* 14 (1986), 597-610.
585. FRISTENSKY, B. XYLEM:UNIX software for genetic database subset manipulation. Macromolecules, Genes and Computers Symposium, Waterville Valley, NH, August 13-18, 1989. (abstract only).
586. FROELICH, N., WILLIAMS, L. C., CASAGRANDE, J. T., AND McMILLAN, M. Computer analysis of protein sequencing data, (in) *Proteins: Structure and Function*, L'Italien J.J., (ed.), (Plenum, NY), pp.455-460, 1987.
587. FROMMEL, C. The apolar surface area of amino acids and its empirical correlation with hydrophobic free energy. *J. Theor. Biol.* 111 (1984), 247-260.
588. FROMMEL, C., AND HOLZHUTTER, H.-G. An estimate of the effect of point mutation and natural selection on the rate of amino acid replacement in proteins. *J. Mol. Evol.* 21 (1985), 233-257.
589. FUCHS, C. On the distribution of the nucleotides in seven completely sequenced DNAs. *Gene* 10 (1980), 371-373.
590. FUJIYOSHI-YONEDA, T., YONEDA, S., KITAMURA, K., AMISAKI, T., IKEDA, K., INOUE, M., AND ISHIDA, T. Adaptability of restrained molecular dynamics for tertiary structure prediction: application to *Crotalus atrox* venom phospholipase A₂. *Protein Engineering* 4(4) (1991), 443-450.
591. FUKAMI-KOBAYASHI, K., AND TATENO, Y. Robustness of maximum likelihood tree estimation against different patterns of base substitutions. *J. Mol. Evol.* 32 (1991), 79-91.
592. FULKERSON, D. R., AND GROSS, O. A. Incidence matrices and interval graphs. *Pacific J. Math.* 15 (1965), 835-856.
593. FURLONG, N. B., AND BECKNER, C. F. Periodicities of dinucleotide self-information values in X174 DNA. *Z. Naturforsch.* 1982. 37:321-325.
594. GABORIAUD, C., BISSERY, V., BENCHETRIT, T., AND MORNON, J. P. Hydrophobic cluster analysis: an efficient new way to compare and analyse amino acid sequences. *FEBS Lett.* 224 (1987), 149-155.
595. GABRIELIAN, A. E., IVANOV, V. S., AND KOZHICH, A. T. On searching for the active sites in proteins and peptide hormones. *CABIOS* 6 (1990), 1-2.
596. GALAS, D. J., EGGERT, M., AND WATERMAN, M. S. Rigorous pattern-recognition methods for DNA sequences. *J. Molecular Biology* 186 (1985), 117-128.
597. GALAT, A. Analysis of dynamics trajectories of DNA and DNA-drug complexes. *CABIOS* 5 (1989), 271-278.
598. GALAT, A. CORGEN: A FORTRAN-77 generator of standard and non-standard DNA helices from the sequence. *CABIOS* 5(4) (1989), 279-286.
599. GALAT, A. A note on graphing helical parameters of dynamics structure of DNA. *J. Mol. Graphics* 8(3) (1990), 173-178.
600. GALIL, Z. Optimal parallel algorithms for string matching, (in) *Proceedings of the ACM Symposium on Theory of Computing*, pp.240-248, 1984.
601. GALIL, Z., AND GIANCARLO, R. On the exact complexity of string matching: upper bounds. *SIAM J. Computing* 21(3) (1992), in press.
602. GALIL, Z., AND GIANCARLO, R. String matching in real time. *J. ACM* 28 (1981), 134-149.
603. GALIL, Z., AND GIANCARLO, R. Improved string matching with k mismatches, *SIGACT News*, April, pp.52-54, 1986.
604. GALIL, Z., AND GIANCARLO, R. Parallel string matching with k mismatches. *TCS* 51 (1987), 341-348.
605. GALIL, Z., AND GIANCARLO, R. Speeding up dynamic programming with applications to molecular biology. Tech. Rep. 110-87, Columbia University Department of Computer Science, 1987.
606. GALIL, Z., AND GIANCARLO, R. Data structures and algorithms for approximate string matching, a survey. *J. Complexity* 4 (1988), 33-72.
607. GALIL, Z., AND GIANCARLO, R. Speeding up dynamic programming with applications to molecular biology. *Theoretical Computer Science* 65 (1989), 107-118.
608. GALIL, Z., AND PARK, K. An improved algorithm for approximate string matching. *SIAM J. Computing* 19 (1990), 989-999.
609. GALIL, Z., AND PARK, K. A linear-time algorithm for concave one-dimensional dynamic programming. *Info. Proc. Letters* 33 (1990), 309-311.
610. GALIL, Z., AND RABANI, Y. On the space requirement of the fastest algorithms for computing edit distance with convex or concave gap costs. *Theoretical Computer Science*, in press.
611. GALIL, Z., AND SEIFERAS, J. Time space optimal string matching. *J. Computer and System Sciences* 26 (1983), 280-294.
612. GALLANT, J., MAIER, D., AND STORER, J. On finding minimal length superstrings. *Journal of Computer and System Science* 20 (1980), 50-58.
613. GALLANT, J. K. The complexity of the overlap method for sequencing biopolymers. *J. Theor. Biol.* 101 (1983), 1-17.

614. GANOZA, M. C., KOFOID, E. C., MARLIERE, P., AND LOUIS, B. G. Potential secondary structure at translation-initiation sites. *Nucl. Acids Res.* 15 (1987), 345-360.
615. GARDEN, P. W. Markov analysis of viral DNA/RNA sequences. *J. Theor. Biol.* 82 (1980), 679-684.
616. GARDINER-GARDEN, M., AND FROMMER, M. CpG islands in vertebrate genomes. *J. Mol. Biol.* 196 (1987), 261-282.
617. GARELL, J. I. Two-dimensional gel electrophoresis and computer analysis of proteins synthesized by clonal cell lines. *J. Biol. Chem.* 254 (1979), 7961-7977.
618. GAREY, M. R., AND JOHNSON, D. S. *Computer and Intractability*, (W.H. Freeman and Company, NY), 1979.
619. GARNIER, J., OSGUTHORPE, D. J., AND ROBSON, B. Analysis of the accuracy and implications of simple methods for predicting the secondary structure of globular proteins. *J. Mol. Biol.* 120 (1978), 97-120.
620. GARRATT, R. C., TAYLOR, W. R., AND THORNTON, J. M. The influence of tertiary structure on secondary structure prediction: accessibility versus predictability for beta-structure. *FEBS Lett.* 188 (1985), 59-62.
621. GASCUEL, O. Structural descriptions, discrimination and learning of these descriptions. *Biochimie* 67 (1985), 499-507.
622. GASCUEL, O., AND GOLMARD, J. L. A simple method for predicting the secondary structure of globular proteins: implications and accuracy. *CABIOS* 4(2) (1988), 357-365.
623. GATES, M. A simple way to look at DNA. *J. Theor. Biol.* 119 (1986), 319-328.
624. GATES, M. A. Simpler DNA sequence representations. *Nature* 316 (1985), 219.
625. GATLIN, L. L. The information content of DNA. *J. Theor. Biol.* 18 (1968), 181-194.
626. GATLIN, L. L. Conservation of shannon's redundancy for proteins. *J. Mol. Evol.* 3 (1974), 189-208.
627. GATTERDAM, R. W. Splicing systems and regularity. *Int. J. Comp. Math.* 31 (1989), 63-67.
628. GAUTHERET, D., MAJOR, F., AND CEDERGREN, R. Computer modeling and display of RNA secondary and tertiary structures, (in) *Methods in Enzymology-Molecular Evolution: Computer Analysis of Protein and Nucleic Acid Sequences*, Vol.183 Doolittle, R.F., (ed.), (Academic Press, NY), pp.318-330, 1990.
629. GAUTHERET, D., MAJOR, F., AND CEDERGREN, R. Pattern searching/alignment with RNA primary and secondary structures: an effective descriptor for tRNA. *CABIOS* 6(4) (1990), 325-332.
630. GAUTIER, C., GOUY, M., AND LOUAIL, S. Non-parametric statistics for nucleic acid sequence study. *Biochimie* 67 (1985), 449-453.
631. GAUTIER, N., AND PAVE, A. Object-centered representation for species systematics and identification in living systems in nature. *CABIOS* 6(4) (1990), 383-386.
632. GELFAND, M. S. Statistical analysis of mammalian pre-mRNA splicing sites. *Nucl. Acids Res.* 17 (1989), 6369-6382.
633. GELFAND, M. S. Statistical analysis of mammalian splicing sites. *Mol. Biol.* 23 (1989), 1138-1147.
634. GELFAND, M. S. Computer prediction of exon-intron structure of mammalian pre-mRNAs. *Nucl. Acids Res.* 18 (1990), 5865-5869.
635. GELFAND, M. S. Global methods for the computer prediction of protein-coding regions in nucleotide sequences. *Biotechnology Software* 7 (1990), 3-11.
636. Genetic Algorithms and their Applications: Proceedings of the Second International Conference on Genetic Algorithms, July 28-31, Massachusetts Institute of Technology, Cambridge, MA. (Lawrence Erlbaum Associates: Hillsdale, NJ).
637. GENTLEMAN, J. F., AND MULLIN, R. C. The distribution of the frequency of occurrence of nucleotide subsequences, based on their overlap capability. *Biometrics* 45 (1989), 35-52.
638. GENTLEMAN, J. F., SHADBOLT-FORBES, M. A., AND HAWKINS, J. W., (eds.). Problems of pattern recognition in nucleotide sequences. *Math. Scientist* 9 (1984), 125-139.
639. Geometry Supercomputing Project Report, University of Minnesota Supercomputing Projec, 1989.
640. GEORGE, D. G., BARKER, W., AND HUNT, L. T. The protein identification resource (PIR). *Nucl. Acids Res.* 14(1) (1986), 11-15.
641. GEORGE, D. G., BARKER, W. C., AND HUNT, L. T. Mutation data matrix and its uses, (in) *Methods in Enzymology-Molecular Evolution: Computer Analysis of Protein and Nucleic Acid Sequences*, Vol.183 Doolittle, R.F., (ed.), (Academic Press, NY), pp.333-351, 1990.
642. GEORGE, D. G., BARKER, W. C., AND T., H. L. The Protein Identification Resource (PIR): an on-line computer system for the characterization of proteins based on comparisons with previously characterized protein sequences, (in) *Proteins: Structure and Function*, L'Italien J.J., (ed.), (Plenum, NY), pp.445-453, 1987.
643. GEORGE, D. G., HUNT, L., AND BARKER, W. C. The National Biomedical Research Foundation protein sequence database, (in) *Computational Molecular Biology. Sources and Methods for Sequence Analysis*, Lesk, A. M., (ed.), (Oxford University Press, Oxford), pp.17-26, 1988.
644. GEORGE, D. G., HUNT, L. T., AND BARKER, W. C. Nucleic acid database management. *Science* 235 (1987), 1562.
645. GEORGE, D. G., MEWES, H. W., AND KIHARA, H. A standardization format for sequence data exchange. *Protein Seq. Data Anal.* 1 (1987), 27-39.

646. GERTNER, B. J., WHITNELL, R. M., AND WILSON, K. B. Fast space-filling molecular graphics using dynamic partitioning among parallel processors. *J. Mol. Graphics* 9(3) (1991), 139-147.
647. GEX-FABRY, M., MARCHETTI, F., NAKATA, K., AND DELISI, C. Identifying weak signals in the presence of noise: a new method of locating potential ligand contact residues in immunoglobulin-related molecules. *Mol. Immunology* 26(12) (1989), 1169-1178.
648. GHOSH, D. A relational database of transcription factors. *Nucl. Acids Res.* 18 (1990), 1749-1756.
649. GIBBONS, A. Evolving similarities - between disciplines. *Science* 250 (1990), 504-506.
650. GIBBONS, A., AND RYTTER, W. *Efficient Parallel Algorithms*, (Cambridge University Press, NY), 1988.
651. GIBBS, A. J., AND MCINTIRE, G. A. The diagram: A method for comparing sequences. *Eur. J. Biochem.* 16 (1970), 1-11.
652. GIBRAT, J.-F., GARNIER, J., AND ROBSON, B. Further developments of protein secondary structure prediction using information theory. New parameters and consideration of residue pairs. *J. Mol. Biol.* 198 (1987), 425-443.
653. GICKUEL-SANZEY, B., AND COSSART, P. Homologies between different prokaryotic DNA-binding regulatory proteins and between their sites of action. *EMBO J.* 1 (1982), 591-595.
654. GILBERT, W. Genome sequencing: Creating a new biology for the twenty-first century. *Issues in Science and Technology* 3 (1987), 26-35.
655. GILL, S. C., AND VON HIPPEL, P. H. Calculation of protein extinction coefficients from amino acid sequence data. *Anal. Biochem.* 182 (1989), 319-326.
656. GILLETT, W., AND HEIDEMANN, J. DNA mapping algorithms: The DNA simulator. Tech. Rep. WUCS-90-37, Washington University in St. Louis, 1990.
657. GILMORE, P. C., AND HOFFMAN, A. J. A characterization of comparability graphs and of interval graphs. *Canadian J. Math.* 16 (1964), 539-548.
658. GILNA, P., TOMLINSON, L. J., AND BURKS, C. Submission of nucleotide sequence data to GenBank. *J. General Microbiology* 135 (1989), 1779-1786.
659. GINGERAS, T., MILAZZO, J. P., AND ROBERTS, R. J. A computer assisted method for the determination of restriction enzyme recognition sites. *Nucl. Acids Res.* 5 (1978), 4105-4127.
660. GINGERAS, T. R. Computers and DNA sequences: A natural combination, (in) *Statistical analysis of DNA sequence data*, Weir, B.S., (Marcel-Dekker, NY), pp.150-43, 1983.
661. GINGERAS, T. R., MILAZZO, J. P., SCIAKY, D., AND ROBERTS, R. J. Computer programs for the assembly of DNA sequences. *Nucl. Acids Res.* 7 (1979), 529-545.
662. GINGERAS, T. R., RICE, P. J., AND ROBERTS, R. J. A semi-automated method for the reading of nucleic acid sequencing gels. *Nucl. Acids Res.* 10 (1982), 103-114.
663. GINGERAS, T. R., AND ROBERTS, R. J. Steps toward computer analysis of nucleotide sequences. *Science* 209 (1980), 1322-1328.
664. GINSBURG, M. Online services, (in) *Nucleic Acid and Protein Sequence Analysis: A Practical Approach*, Bishop, M.J., Rawlings, C.J., (eds.), (IRL Press, Oxford), pp.359-385, 1987.
665. GINSBURG, M. Database searching: a short comparative study, *Imperial Cancer Fund Report*, UK12.10, 1988.
666. GO, N. Theoretical studies of protein folding. *Ann. Rev. Biophys. Bioeng.* 12 (1983), 183-210.
667. GO, N. Protein conformation and gene evolution. *Protein, Nucleic Acid and Enzyme* 29 (1986), 90-100.
668. GO, N., AND ABE, H. Randomness of the process of protein folding. *Int. J. Pept. Protein Res.* 22 (1983), 622-632.
669. GO, N., AND TAKETOMI, H. Respective roles of short- and long-range interactions in protein folding. *Proc. Natl. Acad. Sci. USA* 75 (1978), 559-563.
670. GO, N., AND TAKETOMI, H. Studies on protein folding, unfolding and fluctuations by computer simulation. III. Effect of short-range interactions. *Int. J. Pept. Protein Res.* 13 (1979), 235-252.
671. GO, N., AND TAKETOMI, H. Studies on protein folding, unfolding and fluctuations by computer simulation. IV. Hydrophobic interactions. *Int. J. Pept. Protein Res.* 13 (1979), 447-461.
672. GOAD, W. B. Computational analysis of genetic sequences. *Ann. Rev. Biophys. Biophys. Chem.* 15 (1986), 79-95.
673. GOAD, W. B., AND KANEHISA, M. I. Pattern recognition in nucleic acid sequences. I. A general method for finding local homologies and symmetries. *Nucl. Acids Res.* 10(1) (1982), 247-263.
674. GOEL, N., ROUYANIAN, B., AND SANATI, M. On the computation of the tertiary structure of globular proteins. III. Interresidue distances and computed structures. *J. Theor. Biol.* 99 (1982), 705-757.
675. GOJOBORI, T. MORIYAMA, E. N., AND KIMURA, M. Statistical methods for estimating sequence divergence, (in) *Methods in Enzymology-Molecular Evolution: Computer Analysis of Protein and Nucleic Acid Sequences*, Vol.183 Doolittle, R.F., (ed.), (Academic Press, NY), pp.531-550, 1990.
676. GOJOBORI, T., ISHII, K., AND NEI, M. Estimation of average number of nucleotide substitutions when the rate of substitution varies with nucleotide. *J. Mol. Evol.* 18 (1982), 414-423.
677. GOLDBERG, D. E. *Genetic Algorithms in Search, Optimization, and Machine Learning*, (Addison-Wesley, Reading, MA), 1989.

678. GOLDING, G. B. Estimates of DNA and protein sequence divergence: an examination of some assumptions. *Mol. Biol. Evol.* 1 (1983), 125–142.
679. GOLDING, G. B., AND STROBECK, C. Expected frequencies of codon use as a function of mutation rates and codon fitness. *J. Mol. Evol.* 18 (1982), 379–386.
680. GOLOVANOV, E. I., AND NEDOSPASOV, S. A. An effective approach to the identification of subclones of a DNA of known nucleotide sequence. *CABIOS* 2 (1986), 297–299.
681. GOLUMBIC, M. C. *Algorithmic graph theory and perfect graphs*, (Academic Press, NY), 1980.
682. GONNET, G. H., AND BAEZA-YATES, R. A. An analysis of the Karp-Rabin string matching algorithm. *Info. Proc. Letters* 34 (1990), 271–274.
683. GOODALL, G. J., AND FILIPOWICZ, W. The AU-rich sequences present in the introns of plant nuclear pre-mRNAs are required for splicing. *Cell* 58 (1989), 473–483.
684. GOODMAN, M., MOORE, G. W., BARNABAS, J., AND MATSUDA, G. The phylogeny of human globin genes investigated by the maximum parsimony method. *J. Mol. Evol.* 3 (1974), 1–48.
685. GOODRICH, J. A., SCHWARTZ, M. L., AND MCCLURE, W. R. Searching for and predicting the activity of sites for DNA binding proteins: compilation and analysis of the binding sites for *Escherichia coli* integration host factor. *Nucl. Acids Res.* 18 (1990), 4993–5000.
686. GOODSELL, D. S. RMS: Programs for generating raster molecular surfaces. *J. Mol. Graphics* 6 (1988), 41–44.
687. GOODSELL, D. S., MIAN, S., AND OLSON, A. J. Rendering volumetric data in molecular systems. *J. Mol. Graphics* 7 (1989), 41–47.
688. GOODSELL, D. S., AND OLSON, A. J. Automated docking of substrates to proteins by simulated annealing. *Proteins: Structure, Function, and Genetics* 8 (1990), 195–202.
689. GORDON, A. D. A sequence-comparison statistic and algorithm. *Biometrika* 60 (1973), 197–200.
690. GOTOH, O. An improved algorithm for matching biological sequences. *J. Mol. Biol.* 162 (1982), 705–708.
691. GOTOH, O. Alignment of three biological sequences with an efficient traceback procedure. *J. Theor. Biol.* 121 (1986), 327–337.
692. GOTOH, O. Pattern matching of biological sequences with limited storage. *CABIOS* 3(1) (1987), 17–20.
693. GOTOH, O. Consistency of optimal sequence alignments. *Bull. Math. Biol.* 52 (1990), 509–525.
694. GOTOH, O. Optimal sequence alignment allowing for long gaps. *Bull. Math. Biol.* 52 (1990), 359–373.
695. GOTOH, O., AND TAGASHIRA, Y. Sequence search on a supercomputer. *Nucl. Acids Res.* 14(1) (1986), 57–64.
696. GOUY, M. Codon contexts in enterobacterial and coliphage genes. *Mol. Biol. Evol.* 44 (1987), 426–444.
697. GOUY, M. Secondary structure prediction of RNA, (in) *Nucleic Acid and protein Sequence Analysis: A Practical Approach*, Bishop, M.J., Rawlings, C.J., (eds.), (IRL Press, Oxford), pp.359–385, 1987.
698. GOUY, M., AND GAUTIER, C. Codon usage in bacteria: correlation with gene expressivity. *Nucl. Acids Res.* 10 (1982), 7055–7074.
699. GOUY, M., GAUTIER, C., ATTIMONELLI, M., LANAVE, C., AND di Paola, G. ACNUC—a portable retrieval system for nucleic acid sequence databases: logical and physical designs and usage. *CABIOS* 1(3) (1985), 167–172.
700. GOUY, M., GAUTIER, C., AND MILLERET, F. System analysis and nucleic acid sequence banks. *Biochimie* 67 (1985), 433–436.
701. GOUY, M., MARLIERE, P., PAPANICOLAOU, C., AND NINIO, J. Prediction of secondary structures of nucleic acids: algorithmic and physical aspects. *Biochimie* 67 (1985), 523–531.
702. GOUY, M., MILLERET, F., MUGNIER, C., JACOBZONE, M., AND GAUTIER, C. ACNUC: a nucleic acid sequence data base and analysis system. *Nucl. Acids Res.* 12 (1984), 121–127.
703. GRANIERO-PORATI, M. I., PORATI, A., AND ZANI, L. Informational parameters of an exact DNA base sequence. *J. Theor. Biol.* 86 (1980), 401–403.
704. GRANIERO-PORATI, M. I., PORATI, A., AND ZANI, L. Informational parameters and randomness of mitochondrial DNA. *J. Mol. Evol.* 27 (1988), 109–113.
705. GRANTHAM, R. Amino acid difference formula to help explain protein evolution. *Science* 185 (1974), 862–864.
706. GRANTHAM, R., GAUTIER, C., AND GOUY, M. Codon frequencies in 119 individual genes confirm consistent choices of degenerate bases according to genome type. *Nucl. Acids Res.* 8 (1980), 1893–1912.
707. GRANTHAM, R., GAUTIER, C., GOUY, M., JACOBZONE, M., AND MERCIER, R. Codon catalog usage is a genome strategy modulated for gene expressivity. *Nucl. Acids Res.* 9 (1981), 43–74.
708. GRANTHAM, R., GAUTIER, C., GOUY, M., MERCIER, R., AND PAVE, A. Codon catalog usage and the genome hypothesis. *Nucl. Acids Res.* 8 (1980), r49–r62.
709. GRANTHAM, R., PERRIN, P., AND MOUCHIROUD, D. Patterns of codon usage of different kind of species. *Oxford Surv. Evol. Biol.* 3 (1986), 48–81.
710. GRANUM, E., AND THOMASON, M. G. Automatically inferred Markov network models for classification of chromosomal band pattern structure. *Cytometry* 11 (1990), 26–39.
711. GRAY, A. J., BEECHER, D. E., AND OLSON, M. V. Computer-based image analysis of one-dimensional electrophoretic gels used for the separation of DNA restriction fragments. *Nucl. Acids Res.* 12 (1984), 473–491.

712. GRAY, M. W., SANKOFF, D., AND CEDERGREN, R. J. On the evolutionary descent of organisms and organelles: a global phylogeny based on a highly conserved structural core in small subunit ribosomal RNA. *Nucl. Acids Res.* **12** (1984), 5837-5852.
713. GRAY, N. A program to find regions of similarity between homologous protein sequences using dot-matrix analysis. *J. Mol. Graphics* **8** (1990), 11-15.
714. GREENES, R. A. , (ed.). Symposium On Computer Applications in Medical Care-1988, (IEEE Computer Society Press, Los Angeles), 1989.
715. GREER, J. Protein structure and function by comparative model building. *Ann. NY. Acad. Sci.* **439** (1985), 44-63.
716. GREER, J. Comparative modeling of homologous proteins, (in) *Methods in Enzymology-Molecular design and Modeling: Concepts and Applications*, Part A, Vol.202, Langone, J.J., (ed.), (Academic Press, NY), pp.239-252, 1991.
717. GRIBSKOV, M., DEVEREUX, J., AND BURGESS, R. R. The codon preference plot: Graphic analysis of protein coding sequences and prediction of gene expression. *Nucl. Acids Res.* **12** (1984), 539-549.
718. GRIBSKOV, M., HOMYAK, M., EDENFIELD, J., AND EISENBERG, D. Profile scanning for three-dimensional structural patterns in protein sequences. *CABIOS* **4**(1) (1988), 61-66.
719. GRIBSKOV, M., LUTHY, R., AND EISENBERG, D. Profile analysis, (in) *Methods in Enzymology-Molecular Evolution: Computer Analysis of Protein and Nucleic Acid Sequences*, Vol.183 Doolittle, R.F., (ed.), (Academic Press, NY), pp.146-159, 1990.
720. GRIBSKOV, M., McLACHLAN, A. D., AND EISENBERG, D. Profile analysis: program to find regions of similarity between homologous protein sequences using dot-matrix analysis: Detection of distantly related proteins. *Proc. Natl. Acad. Sci. USA* **84** (1987), 4355-4358.
721. GRIGGS, J. R., AND WATERMAN, M. S. Linked partitions and maps of DNA, 1984. preprint.
722. GRIGORIADOS, M. D., AND KALANTARI, B. A lower bound to the complexity of euclidean and rectilinear matching algorithms. Tech. Rep. LCSR-TR-69, Rutgers University, 1985.
723. GROB, U., AND STUBER, K. Recognition of ill-defined signals in nucleic acid sequences. *CABIOS* **4** (1988), 79-88.
724. GROSJEAN, H., AND FIERS, W. Preferential codon usage in prokaryotic genes: the optimal codon-anticodon interaction energy and the selective codon usage in highly expressed genes. *Gene* **18** (1982), 199-209.
725. GROSJEAN, H. J., SANKOFF, D., MIN JOU, W., FIERS, W., AND CEDERGEREN, R. J. Bacteriophage MS2 RNA: a correlation between the stability of codon-anticodon adaptation and the choice of codon words. *J. Mol. Evol.* **12** (1979), 113-119.
726. GROSS, E., AND DUANE, J. Protein graphics: historical development, future directions and microcomputer applications. *CABIOS* **2** (1986), 173-179.
727. GROSSI, R., AND LUCCIO, R. Simple and efficient string matching with k mismatches. *Info. Proc. Letters* **33** (1989), 113-120.
728. GRUSKIN, K. D., AND SMITH, T. F. Molecular genetics and computer analyses. *CABIOS* **3**(3) (1987), 167-170.
729. GRYMES, R. A., TRAVERS, P., AND ENGELBERG, A. Gel - a computer tool for DNA sequencing projects. *Nucl. Acids Res.* **14** (1986), 87-99.
730. GUIBAS, L. J., AND ODLYZKO, A. M. Long repetitive patterns in random sequences. *Z. Wahrscheinlichkeitstheorie verw. Gebiete* **53** (1980), 241-262.
731. GUIBAS, L. J., AND ODLYZKO, A. M. Periods in strings. *J. Combin. Theory Ser. A* **30** (1981), 19-42.
732. GUIBAS, L. J., AND ODLYZKO, A. M. String overlaps, pattern matching and nontransitive games. *J. Combin. Theory Ser. A* **30** (1981), 183-208.
733. GUIGO, R., JOHANSSON, A., AND SMITH, T. F. Automatic evaluation of protein sequence functional patterns, preprint, 1989.
734. GUIGO, R., JOHANSSON, A., AND SMITH, T. F. Automatic evaluation of protein sequence functional patterns. *CABIOS* **7**(3) (1991), 309-315.
735. GULTYAEV, A. P. The computer simulation of RNA folding involving pseudoknot formation. *Nucl. Acids Res.* **19**(9) (1991), 2489-2494.
736. GUNTERT, P., AND BRAUN, W. WUTHRICH, K. Efficient computation of three-dimensional protein structures in solution from nuclear magnetic resonance data using the program DIANA and the supporting programs CALIBA, HABAS and GLOMSA. *J. Mol. Biol.* **217** (1991), 517-530.
737. GUSEIN-ZADE, S. M., AND BORODOVSKY, M. Y. An improved distribution of codon frequencies allowing for inhomogeneity of DNA's primary-structure evolution. *J. Biomol. Struct. Dyn.* **7** (1990), 1185-1197.
738. GUSTAFSON, J. L., MONTRY, G. R., AND BENNER, R. E. Development of parallel methods for a 1024-processor hypercube. *SIAM J. Scientific and Statistical Computing* **9** (1988), 609.
739. GUTMAN, G. A., AND HATFIELD, G. W. Nonrandom utilization of codon pairs in *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* **86** (1990), 3699-3703.
740. GUY, H. R. Amino acid side-chain partition energies and distribution of residues insoluble proteins. *Biophys. J.* **47** (1985), 61-70.
741. GWILLIAM, M., AND MAX, N. Atoms with shadows — an area based algorithm for cast shadows on space-filling molecular models. *J. Mol. Graphics* **7** (1989), 54-59.

742. HABER, J. E., AND KOSHLAND, D. E. Jr. An evaluation of the relatedness of proteins based on comparison of amino acid sequences. *J. Mol. Biol.* **50** (1970), 617-639.
743. HAGERUP, T., AND RUB, C. Optimal merging and sorting on the EREW PRAM. *Info. Proc. Letters* **33** (1989), 181-186.
744. HAGLER, A., AND HONIG, B. On the formation of protein tertiary structure on a computer. *Proc. Natl. Acad. Sci. USA* **75** (1978), 554-558.
745. HAHN, S., BURATOWSKI, S., SHARP, P. A., AND GUARENTE, L. Isolation of the gene encoding the yeast TATA binding protein TFIID: A gene identical to the SPT15 suppressor of ty element insertions. *Cell* **58** (1989), 1173-1181.
746. HAIECH, J., AND SALLANTIN, J. Computer search of calcium binding sites in a gene data bank: use of learning techniques to build an expert system. *Biochimie* **67** (1985), 555-560.
747. HALL, J., AND MYERS, E. A software tool for finding locally optimal alignments in protein and nucleic acid sequences. *CABIOS* **4**(1) (1988), 35-40.
748. HALL, M. Searching the ultimate software. *Sun Technology*, Summer (1989), 41-49.
749. HALL, S. S. James Watson and the search for biology's "Holy Grail", 1990. Feb. 40-49.
750. HAMILTON, R., WATANABE, C. K., AND DE BOER, H. A. Compilation and comparison of the sequence context around the AUG start codons in *Saccharomyces cerevisiae* mRNAs. *Nucl. Acids Res.* **15** (1987), 3581-3594.
751. HAMM, G. H., AND CAMERON, G. N. The EMBL data library. *Nucl. Acids Res.* **14** (1986), 5-9.
752. HAMODRAKAS, S. J. A protein secondary structure prediction scheme for the IBM PC and compatibles. *CABIOS* **4** (1988), 473-477.
753. HAMORI, E. Computer-based abstract representation of very long nucleotide sequences. *Fed. Proc.* **40** (1981), 1647.
754. HAMORI, E. Low resolution h curve of the 48,502 nucleotides long lambda genome. *Gene Anal. Techn.* **1** (1984), 69-74.
755. HAMORI, E. Novel DNA sequence representations. *Nature* **314** (1985), 585-586.
756. HAMORI, E. Long range nucleotide comparison patterns in fully sequenced genomes. *Biol. Chem. Hoppe-Seyler* **367S** (1986), 226.
757. HAMORI, E. Graphic representation of long DNA sequences by the methods of H curves—Current results and future aspects. *BioTechniques* **7**(7) (1989), 710-720.
758. HAMORI, E. Compact graphic representation of very long DNA sequences: Plans for communicating these over computer network, (in) *Proc. of the First Int. Conf., on Electrophoresis, Supercomputing and the Human Genome*, Cantor C.R., Lim H.A., (eds.), (World Scientific, NJ), 1991.
759. HAMORI, E., AND RUSHKIN, J. H Curves—A novel method of representation of nucleotide series especially suited for long DNA sequences. *J. Biol. Chem.* **258** (1983), 1318-1327.
760. HAMORI, E., AND VARGA, G. DNA sequence (H) curves of the human immunodeficiency virus i and some related virus genomes. *DNA* **7** (1988), 371-378.
761. HAMORI, E., AND VARGA, G. Use of H curves in searches for DNA sequence which code for overlapping peptide chains. *FASEB J.* **3** (1988), A331.
762. HAMORI, E., VARGA, G., AND LAGUARDIA, J. J. HYLAS: program for generating H curves (abstract 3-D representations of long DNA sequences). *CABIOS* **5**(4) (1989), 263-269.
763. HANAI, R., AND WADA, A. The effects of guanine and cytosine variation on dinucleotide frequency and amino acid composition in the human genome. *J. Mol. Evol.* **27** (1988), 321-325.
764. HANAI, R., AND WADA, A. Novel third-letter bias in *Escherichia coli* codons revealed by rigorous treatment of coding constraints location. *J. Mol. Biol.* **207** (1989), 655-660.
765. HANAI, R., AND WADA, A. Doublet preference and gene evolution. *J. Mol. Evol.* **30** (1990), 109-115.
766. HANASH, S. M., STRAHLER, J. R., NEEL, J. V., HAILAT, N., MELHEM, R., KEIM, D., ZHU, X. X., WAGNER, D., GAGE, D. A., AND WATSON, J. T. Highly resolving two-dimensional gels for protein sequencing. *Proc. Natl. Acad. Sci. USA* **88** (1991), 5709-5713.
767. HANLEY, B. A., AND SCHULER, M. A. Plant intron sequences: evidence for distinct groups of introns. *Nucl. Acids Res.* **16** (1988), 7159-7176.
768. HAO, M.-H., AND OLSON, W. K. Global equilibrium configurations of supercoiled DNA. *Macromolecules* **22** (1989), 3292-3303.
769. HAO, M.-H., AND OLSON, W. K. Modeling DNA supercoils and knots with B-spline functions. *Biopolymers* **28** (1989), 873-900.
770. HAO, M.-H., AND OLSON, W. K. Molecular modeling and energy refinement of supercoiled DNA. *J. Biomol. Struct. Dyn.* **7**(3) (1989), 661-692.
771. HAREL, D., AND TARJAN, R. E. Fast algorithms for finding nearest common ancestors. *SIAM J. Computing* **13** (1984), 338-355.
772. HARLEY, C. B., AND REYNOLDS, R. P. Analysis of *E. coli* promoter sequences. *Nucl. Acids Res.* **15** (1987), 2343-2361.
773. HARPER, D. S., AND JAHN, C. L. Differential use of termination codons in higher eukaryotes. *Proc. Natl. Acad. Sci. USA* **86** (1989), 3252-3256.

774. HARR, R., HAGBLOM, P., AND GUSTAFSSON, P. Two-dimensional graphic analysis of DNA sequence homologies. *Nucl. Acids Res.* **10**(1) (1982), 363-374.
775. HARR, R., HAGGSTROM, M., AND GUSTAFSSON, P. Search algorithm for pattern match analysis of nucleic acid sequences. *Nucl. Acids Res.* **11**(9) (1983), 2943-2957.
776. HARRIS, L. F., SULLIVAN, M. R., AND HICKOK, D. Conservation of genetic information between regulatory protein DNA binding alpha helices and their cognate operator sites: A simple code for site-specific recognition. *Computers Math. Appl.* **20** (1990), 1-23.
777. HARRIS, L. F., SULLIVAN, M. R., AND HICKOK, D. Genetic sequences of hormone response elements share similarity with predicted alpha helices within DNA binding domains of steroid receptor proteins: A basis for site-specific recognition. *Computers Math. Appl.* **20** (1990), 25-48.
778. HARRIS, N. L., AND SENAPATHY, P. Distribution and consensus of branch point signals in eukaryotic genes: a computerized statistical analysis. *Nucl. Acids Res.* **18** (1990), 3015-3019.
779. HARTWIG, M., MATTHEWS, E., AND ARNOLD, W. Extremely underwound chromosomal DNA in nucleoids of mouse sarcoma cells. *Cancer Lett.* **13** (1981), 153.
780. HASEGAWA, M., YASUNAGA, T., AND MIYATA, T. Secondary structure of MS2 phage RNA and bias in code word usage. *Nucl. Acids Res.* **7** (1979), 2073-2079.
781. HASEGAWA, M., KISHINO, H., AND SAITOU, N. On the maximum likelihood method in molecular phylogenetics. *J. Mol. Evol.* **32** (1991), 443-445.
782. HASTINGS, K., AND EMERSON, C. Codon usage in muscle genes and liver genes. *J. Mol. Evol.* **19** (1983), 214-218.
783. HATADA, I., HAYASHIZAKI, Y., HIROTSUNE, S., KOMATSUBARA, H., AND MUKAI, T. A genomic scanning method for higher organisms using restriction sites as landmarks. *Proc. Natl. Acad. Sci. USA* **88** (1991), 9523-9527.
784. HATFIELD, D., AND EMERSON, C. P. Aminoacyl-tRNA (anticodon):codon adaptation in human and rabbit reticulocytes. *Biochem. Int.* **13** (1986), 835-842.
785. HAVEL, T., CRIPPEN, G. M., AND KUNTZ, I. D. Effects of distance constraints on macromolecular conformation. II. Simulation of experimental results and theoretical predictions. *Biopolymers* **18** (1979), 73-81.
786. HAVEL, T., KUNTZ, I. D., AND CRIPPEN, G. M. Theory and practice of distance geometry. *Bull. Math. Biol.* **45** (1983), 665-720.
787. HAVEL, T., AND WUTHRICH, K. A distance geometry program for determining the structures of small proteins and other macromolecules for nuclear magnetic resonance measurements of intramolecular 1H - 1H proximities in solution. *Bull. Math. Biol.* **46** (1984), 673-698.
788. HAWKINS, J. D. A survey on intron and exon lengths. *Nucl. Acids Res.* **16** (1988), 9893-9905.
789. HAWLEY, D. K., AND MCCLURE, W. R. Compilation and analysis of *Escherichia coli* promoter DNA sequences. *Nucl. Acids Res.* **11** (1983), 2237-2255.
790. HAYASHI, K., AND MUNAKATA, N. Basically musical. *Nature* **3110** (1984), 96-99.
791. HEAD, T. Formal language theory and DNA: An analysis of the generative capacity of specific recombinant behaviors. *Bull. Math. Biol.* **49** (1987), 737-759.
792. HEIN, J. A new method that simultaneously aligns and reconstructs ancestral sequences for any number of homologous sequences, when the phylogeny is given. *Mol. Biol. Evol.* **6** (1989), 649-668.
793. HEIN, J. A tree reconstruction method that is economical in the number of pairwise comparisons used. *Mol. Biol. Evol.* **6** (1989), 669-684.
794. HEIN, J. Unified approach to alignment and phylogenies, (in) *Methods in Enzymology-Molecular Evolution: Computer Analysis of Protein and Nucleic Acid Sequences*, Vol.183, Doolittle, R.F., (ed.), (Academic Press, NY), pp.626-645, 1990.
795. HEIN, J. J. An optimal algorithm to reconstruct trees from additive distance data. *Bull. Math. Biol.* **51** (1989), 597-603.
796. HELLINGA, H. W., AND RICHARDS, F. M. Construction of new ligand binding sites in proteins of known structure I. Computer-aided modeling of sites with pre-defined geometry. *J. Mol. Biol.* **222** (1991), 763-785.
797. HENAUT, A., LIMAIEM, J., AND VIGIER, P. The origins of the strategy of codon usage. *Biochimie* **67** (1985), 475-483.
798. HENDY, M. D., AND PENNY, D. Branch and bound algorithms to determine minimal evolutionary trees. *Math. Biosci.* **59** (1982), 277-290.
799. HENIKOFF, S., AND HENIKOFF, J. G. Automated assembly of protein blocks for database searching. *Nucl. Acids Res.* **19**(23) (1991), 6565-6572.
800. HENIKOFF, S., WALLACE, J. C., AND BROWN, J. P. Finding protein similarities with nucleotide sequence databases, (in) *Methods in Enzymology-Molecular Evolution: Computer Analysis of Protein and Nucleic Acid Sequences*, Vol.183, Doolittle, R.F., (ed.), (Academic Press, NY), pp.111-132, 1990.
801. HENNEKE, C. M. A multiple sequence alignment algorithm for homologous proteins using secondary structure information and optionally keying alignments to functionally important sites. *CABIOS* **5**(2) (1989), 141-150.
802. HIGGINS, D. G. Sequence ordinations: a multivariate analysis approach to analysing large sequence data sets. *CABIOS* **8**(1) (1992), 15-22.

803. HIGGINS, D. G., AND SHARP, P. M. Fast and sensitive multiple sequence alignments on a microcomputer. *CABIOS* 5(2) (1989), 151–153.
804. HINDS, P. W., AND BLAKE, R. D. Delineation of coding areas in DNA sequences through assignment of codon probabilities. *J. Biomol. Struct. Dyn.* 3 (1985), 543–549.
805. HIRSCHBERG, D. S. A linear space algorithm for computing longest common subsequences. *CACM* 18 (1975), 341–343.
806. HIRSCHBERG, D. S. Fast parallel sorting algorithms. *CACM* 21 (1978), 657–661.
807. HIRST, J. D., AND STERNBERG, M. J. E. Prediction of ATP-binding motifs: a comparison of a perceptron-type neural network and a consensus sequence method. *Protein Engineering* 4(6) (1991), 615–623.
808. HO, S. T. S., ALLISON, L., AND YEE, C. N. Restriction site mapping for three or more enzymes. *CABIOS* 6(3) (1990), 195–204.
809. HOCKNEY, R. W., AND JESSHOPE, C. R. *Parallel Computers* 2 (Adam Hilger, NY), 1988.
810. HODGMAN, T. C. The elucidation of protein function from its amino acid sequence. *CABIOS* 2(3) (1986), 181–187.
811. HODGMAN, T. C. The elucidation of protein function by sequence motif analysis. *CABIOS* 5(1) (1989), 1–13.
812. HODGSON, C. P. Cloning vector artifacts in the DNA database. *Bio Techniques* 9 (1990), 54.
813. HOGEWEG, P., AND HESPER, B. The alignment of sets of sequences and the construction of phylogenetic trees. An integrated method. *J. Mol. Evol.* 20 (1984), 175–186.
814. HOGEWEG, P., AND HESPER, B. Energy directed folding of RNA sequences. *Nucl. Acids Res.* 12 (1984), 67–74.
815. HOGLUND, M., SALL, T., AND ROHME, D. On the origin of coding sequences from random open reading frames. *J. Mol. Evol.* 30 (1990), 104–108.
816. HOLDEN, C. An omnifarious data bank for biology? *Science* 228 (1985), 1412–1413.
817. HOLLEY, L., AND KARPLUS, M. Protein secondary structure prediction for a neural network. *Proc. Natl. Acad. Sci.* 86 (1989), 152–156.
818. HOLLEY, L. H., AND KARPLUS, M. Protein secondary structure prediction with a neural network. *Proc. Natl. Acad. Sci. USA* 86 (1989), 152–156.
819. HOLLEY, L. H., AND KARPLUS, M. Neural networks for protein structure prediction, (in) *Methods in Enzymology-Molecular design and Modeling: Concepts and Applications*, Part A, Vol. 202, Langone, J. J., (ed.), (Academic Press, NY), pp. 204–224, 1991.
820. HOLM, L. Codon usage and gene expression. *Nucl. Acids Res.* 14 (1986), 3075–3087.
821. HOLM, L., AND SANDER, C. Database algorithm for generating protein backbone and side-chain coordinates from a C^α trace application to model building and detection of co-ordinate errors. *J. Mol. Biol.* 218 (1991), 183–194.
822. HOLMQUIST, G. P. Evolution of chromosome bands: Molecular ecology of noncoding DNA. *J. Mol. Evol.* 28 (1988), 469–486.
823. HOLMQUIST, R. Evaluation of compositional nonrandomness in proteins. *J. Mol. Evol.* 11 (1978), 349–360.
824. HOLMQUIST, R. A measure of the denseness of a phylogenetic network. *J. Mol. Evol.* 11 (1978), 225–231.
825. HOLMQUIST, R. The augmentation algorithm and molecular phylogenetic trees. *J. Mol. Evol.* 12 (1978), 17–24.
826. HOLMQUIST, R. The Reh theory of protein and nucleic acid divergence: a retrospective update. *J. Mol. Evol.* 11 (1978), 361–374.
827. HOLMQUIST, R. The method of parsimony: an experimental test and theoretical analysis of the adequacy of molecular restoration studies. *J. Mol. Biol.* 135 (1979), 939–958.
828. HOLMQUIST, R. Molecular phylogenetic trees: on the validity of the Goodman-Moore augmentation algorithm. *J. Mol. Evol.* 13 (1979), 173–178.
829. HOLMQUIST, R., CANTOR, C. R., AND JUKES, T. H. Improved procedures for comparing homologous sequences in molecules of proteins and nucleic acids. *J. Mol. Biol.* 64 (1972), 145–161.
830. HOLMQUIST, R., AND CONROY, T. The estimation of genetic divergence. *J. Mol. Biol.* 17 (1981), 167–181.
831. HOLMQUIST, R., GOODMAN, M., CONROY, T., AND CZELUSNIAK, J. The spatial distribution of fixed mutations within genes coding for proteins. *J. Mol. Evol.* 19 (1983), 437–448.
832. HOOD, L., AND SMITH, L. Genome sequencing: How to proceed. *Issues in Science and Technology* 3 (1987), 36–46.
833. HOPP, T. P. Identification of proteins surfaces and interaction sites by hydrophilicity analysis, (in) *Proteins: Structure and Function*, L'Italien J. J., (ed.), (Plenum, NY), pp. 437–443, 1987.
834. HOPP, T. P., AND WOODS, K. R. Prediction of protein antigenic determinants from amino acid sequences. *Proc. Natl. Acad. Sci.* 78 (1981), 3824–3828.
835. HORIKOSHI, M., WANG, C. K., FUJII, H., CROMLISH, J. A., WEIL, P. A., AND ROEDER, R. G. Cloning and structure of a yeast gene encoding a general transcription initiation factor TFIID that binds to the TATA box. *Nature* 341 (1989), 299–303.
836. HORIMOTO, K., SUZUKI, H., AND OTSUKA, J. Principal component analysis to detect the similarity of distantly related proteins; its application to cytochromes *c, c₁* and *f*. *Protein Seq. Data Anal.* 4 (1991), 33–42.

837. HOWARTH, A. J., AND VANDEMARK, G. J. Phylogeny of geminiviruses. *J. Gen. Virol.* **70** (1989), 2717-2727.
838. HOWLETT, R. Between biology and culture. *Nature* **347** (1990), 621-622.
839. HOZIER, J., NUTTALL, T., NEWMAN, H., AND SHOAFF, W. Supercomputer assembly and analysis of a digitally simulated human chromosome, (in) *Proc. of the First Int. Conf., on Electrophoresis, Supercomputing and the Human Genome*, Cantor C.R., Lim H.A., (eds.), (World Scientific, NJ), 1991.
840. HUANG, X. A lower bound for the edit-distance problem under an arbitrary cost function. *Info. Proc. Letters* **27** (1988), 319-321.
841. HUANG, X. A space-efficient parallel sequence comparison algorithm for a message-passing multiprocessor. *Int. J. Parallel Programming* **18** (1989), 223-239.
842. HUBBARD, J. M., AND HEARST, J. E. Predicting the three-dimensional folding of transfer RNA with a computer modeling protocol. *Biochemistry* **30** (1991), 5458-5465.
843. HUBBARD, T. J. P., AND BLUNDELL, T. L. Comparison of solvent inaccessible cores of homologous proteins: definitions useful for protein modelling. *Protein Engineering* **1** (1987), 159-171.
844. HUME, A., AND SUNDAY, D. Fast string searching. Tech. Rep. Computing Science Technical Report Number 156, AT&T Bell Labs, 1991.
845. HUNT, L. T. Usefulness of the PIR database for protein comparisons, (in) *Methods in Protein Sequence Analysis*, Jornvall, H., Hoog, J.-O., and Gustavsson, A.-M., (ed.), (Birkhauser Verlag, Boston, MA), pp.343-352, 1991.
846. HUNT, L. T., GEORGE, D. G., AND BARKER, W. C. Short segment searches for compositional similarities in proteins. *Fed. Proc.* **46** (1987), 2231.
847. HURLEY, C. B., AND OLDFORD, R. W. A software model for statistical graphics, pp. 1-15 (manuscript in preparation).
848. HUTCHINSON, G. Evaluation of polymer sequence data from two complete digests. *Internal Report NIH* (1968), 1-73.
849. HUTCHINSON, G. Evaluation of polymer sequence fragment data using graph theory. *Bull. Math. Biophys.* **31** (1969), 541-562.
850. HYDE, J. E., AND SIMS, P. F. G. Anomalous dinucleotide frequencies in both coding and non-coding regions from the genome of the human malaria parasite plasmodium falciparum. *Gene* **61** (1987), 177-178.
851. IBARRA, O. H., PONG, T., AND SOHN, S. M. Hypercube algorithms for some string comparison problem, (in) *Proceedings of the 1988 International Conference on Parallel Processing*, Vol.3, pp. 190-193, 1988.
852. IBARRA, O. H., PONG, T., AND SOHN, S. M. String processing on the hypercube. *IEEE Transactions on Acoustics, Speech, and Signal Processing* **38** (1990), 160-164.
853. IIDA, Y. Splice-site signals of mRNA precursors as revealed by computer search. Site-specific mutagenesis and thalassemia. *J. Biochem.* **97** (1985), 1173-1179.
854. IIDA, Y. DNA sequences and multivariate statistical analysis. Categorical discrimination approach to 5' splice site signals of mRNA precursors in higher eukaryotes' genes. *Comput. Appl. Biosci.* **3** (1987), 93-98.
855. IIDA, Y. Categorical discriminant analysis of 3'-splice site signals of mRNA precursors in higher eukaryote genes. *J. Theor. Biol.* **135** (1988), 109-118.
856. IIDA, Y. Quantification analysis of 5'-splice site signal sequences in mRNA precursors. Mutations in rabbit beta-globin gene. *Biochim. Biophys. Acta* **100** (1989), 270-276.
857. IIDA, Y. Quantification analysis of 5'-splice site signal sequences in mRNA precursors. Mutations in 5'-splice signal sequence of human beta-globin gene and beta-thalassemia. *J. Theor. Biol.* **145** (1990), 523-534.
858. IIDA, Y., AND SASAKI, F. Recognition patterns for exon-intron junctions in higher organisms as revealed by computer search. *J. Biochem.* **94** (1983), 1731-1738.
859. IKEMURA, T. The frequency of codon usage in *E.coli* genes: correlation with the abundance of cognate tRNA, Osawa S., Ozeki H., Uchida H., Yura T., (University of Tokyo Press), 1980.
860. IKEMURA, T. Correlation between the abundance of *Escherichia coli* transfer RNAs and the occurrence of the repetitive codon in protein genes. *J. Mol. Biol.* **146** (1981), 1-21.
861. IKEMURA, T. Correlation between the abundance of *escherichia coli* transfer RNAs and the occurrence of the respective codons in its protein genes: A proposal for a synonymous codon choice, that is optimal for the *e.coli* translational system. *J. Mol. Biol.* **151** (1981), 389-409.
862. IKEMURA, T. Correlation between the abundance of yeast transfer RNAs and the occurrence of the repetitive codon in protein genes: differences in synonymous codon choice patterns of yeast and *escherichia coli* with references to the abundance of isoacceptor transfer RNAs. *J. Mol. Biol.* **158** (1982), 573-597.
863. IKEMURA, T. Codon usage and tRNA content in unicellular and multicellular organisms. *Mol. Biol. Evol.* **2** (1985), 13-34.
864. IKEMURA, T., AND AOTA, S. Global variation in G+C content along vertebrate genome DNA: Possible correlation with chromosome band structures. *J. Mol. Biol.* **203** (1988), 2-13.
865. IKEMURA, T., AND OZEKI, H. Codon usage and transfer RNA contents: organism-specific codon-choice patterns in reference to the isoacceptor contents. *Cold Spring Harbor Symp. Quant. Biol.* **47** (1983), 1087-1097.

866. IKEMURA, T., WADA, K., AND AOTA, S. Giant G+C% mosaic structures of the human genome found by arrangement of genbank human DNA segments according to genetic positions. *Genomics* 8 (1990), 207-216.
867. Impulse, National Biomedical Simulation Resource Newsletter, 1989.
868. ISENMAN, M. E., AND SHASHA, D. E. Performance and architectural issues for string matching. *IEEE Transactions on Computers* 39 (1990), 238-250.
869. ISLAM, S. A., AND STERNBERG, M. J. E. A relational database of protein structures designed for flexible enquiries about conformation. *Protein Engineering* 2 (1989), 431-442.
870. Information Processing Systems - Open Systems Interconnection - Specification of Abstract Syntax Notation One (ASN.1), International Organization for Standardization, ISO-8824, pp.1-50, Switzerland , 1987.
871. Information Processing Systems - Open Systems Interconnection - Specification of Basic Encoding Rules for Abstract Syntax Notation One (ASN.1), International Organization for Standardization, ISO-8825, pp.1-15, Switzerland, 1987.
872. IUPAC-IUB Commission on Biochemical Nomenclature. 1966. Abbreviated designation of amino acid derivatives and peptides, tentative rules. *J. Biol. Chem.* 241:2491-2495.
873. IUPAC-IUB Commission on Biochemical Nomenclature. 1968. A one-letter notation for amino acid sequences: tentative rules. *J. Biol. Chem.* 243:3557-3559.
874. IYENGAR, A. K. Parallel characteristics of sequence alignment algorithms. *J. ACM* 11 (1989), 304-313.
875. JACOB, M., AND GALLINARO, H. The 5' splice site: phylogenetic evolution and variable geometry of association with U1RNA. *Nucl. Acids Res.* 17 (1989), 2159-2180.
876. JACOBSON, A. B., GOOD, L., SIMONETTI, J., AND ZUKER, M. Some simple computational methods to improve the folding of large RNAs. *Nucl. Acids Res.* 12 (1984), 45-52.
877. JACOBSON, A. B., KUMAR, H., AND ZUKER, M. Effect of spermidine on the conformation of bacteriophage MS2 RNA. Electron microscopy and computer modeling. *J. Mol. Biol.* 181 (1985), 517-531.
878. JAEGER, J. A., TURNER, D. H., AND ZUKER, M. Improved predictions of secondary structures for RNA. *Proc. Natl. Acad. Sci. USA* 86 (1989), 7706-7710.
879. JAEGER, J. A., TURNER, D. H., AND ZUKER, M. Predicting optimal and suboptimal secondary structure for RNA, (in) *Methods in Enzymology-Molecular Evolution: Computer Analysis of Protein and Nucleic Acid Sequences*, Vol.183, Doolittle, R.F., (ed.), (Academic Press, NY), pp.281-306, 1990.
880. JAENICKE, R. Protein folding: Local structures, domains and assemblies, (in) *Methods in Protein Sequence Analysis*, Jornvall, H., Hoog, J.-O., and Gustavsson, A.-M., (ed.), (Birkhauser Verlag, Boston, MA), pp.387-396, 1991.
881. JAGADEESWARAN, P., AND MCGUIRE, P. M. Jr. Interactive computer programs in sequence data analysis. *Nucl. Acids Res.* 10(1) (1982), 433-447.
882. JAHNIG, F. Structure predictions of membrane proteins are not that bad. *Trends Biochem. Sci.* 15 (1990), 93-95.
883. JAMES, B. D., OLSEN, G. J., AND PACE, N. R. Phylogenetic comparative analysis of RNA secondary structure, (in) *Methods in Enzymology*, Vol.180, (Academic Press, NY), pp.227-239, 1989.
884. JAMESON, B. A. Modelling in peptide design. *Nature* 341 (1989), 465-466.
885. JAMESON, B. A., AND WOLF, H. The antigenic index: a novel algorithm for predicting antigenic determinants. *CABIOS* 4 (1988), 181-186.
886. JANIN, J. Surface and inside volumes in globular proteins. *Nature* 277 (1979), 491-492.
887. Janse Van Rensburg, E. J., AND LIM, H. A. A numerical study of gel electrophoresis of circular DNA, FSU-SCRI-91-41, March 1991.
888. JEFFREY, H. J. Chaos game representation of gene structure. *Nucl. Acids Res.* 18 (1990), 2163-2170.
889. JEMMERSON, R., AND PATERSON, Y. Mapping antigenic sites on proteins: implication for the design of synthetic vaccines. *Biotechniques* 4 (1986), 18-31.
890. JENNINGS, P. A., SAALAU-BETHELL, S. M., FINN, B. E., CHEN, X., AND MATTHEWS, C. B. Mutational analysis of protein folding mechanisms, (in) *Methods in Enzymology-Molecular Design and Modeling: Concepts and Applications*, Part A, Vol.202, Langone, J.J., (ed.), (Academic Press, NY), pp.113-126, 1991.
891. JIANG, K., ZHENG, J., HIGGINS, S. B., WATTERSON, D. M., CRAIG, T. A., LUKAS, T. J., AND Van Eldik, L. J. A knowledge-based experimental design system for nucleic acid engineering. *CABIOS* 6(3) (1990), 205-212.
892. JIMENEZ-MONTANO, M. A. On the syntactic structure of protein sequences and the concept of grammar complexity. *Bull. Math. Biol.* 46 (1984), 641-659.
893. JIN, L., AND NEI, M. Limitations of the evolutionary parsimony method of phylogenetic analysis. *Mol. Biol. Evol.* 7 (1990), 82-102.
894. JIN, L., AND NEI, M. Relative efficiencies of the maximum-parsimony and distance-matrix methods of phylogeny construction for restriction data. *Mol. Biol. Evol.* 8(3) (1991), 356-365.
895. JOHNSON, A. M. Comparison of dinucleotide frequency and codon usage in toxoplasma and plasmodium: evolutionary implications. *J. Mol. Evol.* 30 (1990), 383-387.
896. JOHNSON, L. M. End-user searching of biotech databases. *Bio/Technology* 7 (1989), 378-379.

897. JOHNSON, M. S., AND DOOLITTLE, R. F. A method for the simultaneous alignment of three or more amino acid sequences. *J. Mol. Evol.* **23** (1986), 267-278.
898. JOHNSON, M. S., SALI, A., AND BLUNDELL, T. L. Phylogenetic relationships from three-dimensional protein structures, (in) *Methods in Enzymology-Molecular Evolution: Computer Analysis of Protein and Nucleic Acid Sequences*, Vol.183, Doolittle, R.F., (ed.), (Academic Press, NY), pp.670-690, 1990.
899. JOHNSON, M. S., SUTCLIFFE, M. J., AND BLUNDELL, T. L. Molecular anatomy: phyletic relationships derived from three-dimensional structures of proteins. *J. Mol. Evol.* **30** (1990), 43-59.
900. JOHNSTON, R. E., MACKENZIE, J. M. Jr., AND DOUGHERTY, W. G. Assembly of overlapping DNA sequences by a program written in BASIC for 64k CP/M and MS-DOS IBM-compatible microcomputers. *Nucl. Acids Res.* **14** (1986), 517-527.
901. JONES, D. D. Amino acid properties and side-chain orientations in proteins: a cross correlation approach. *J. Theor. Biol.* **50** (1975), 167-183.
902. JONES, L. P. Portrep: A portable repeated string finder. *Software Practice and Experience* **10** (1980), 63-77.
903. JONES, M. D. DNA sequence analysis and the computer. *Biochem. Soc. Trans.* **12** (1984), 1018-1020.
904. JONES, R., TAYLOR, W., ZHANG, X., MESIROV, J. P., AND LANDER, E. Protein Sequence Comparison on the Connection Machine CM-2, (in) *Computers and DNA, SFI Studies in the Sciences of Complexity*. Vol. VI, Bell, G., Marr, T., (eds.), (Addison-Wesley, NY), 1990.
905. JONES, T. A. A graphics model building and refinement system for macromolecules. *J. Appl. Cryst.* **11** (1978), 268-272.
906. JONES, T. A. Using known substructures in protein model building and crystallography. *EMBO J.* **5** (1986), 819-822.
907. JOSHI, C. P. Putative polyadenylation signals in nuclear genes of higher plants: a compilation and analysis. *Nucl. Acids Res.* **15** (1987), 9627-9640.
908. JUKES, T. H. Codons and nearest-neighbor nucleotide pairs in mammalian messenger RNA. *J. Mol. Evol.* **11** (1978), 121-127.
909. JUKES, T. H. Evolution of the amino acid codes: inferences from mitochondrial codes. *J. Mol. Evol.* **19** (1983), 219-225.
910. JUKES, T. H., AND BHUSHAN, V. Silent nucleotide substitutions and G+C content of some mitochondrial and bacterial genes. *J. Mol. Evol.* **24** (1986), 39-44.
911. JUKES, T. H., AND KIMURA, M. Evolutionary constraints and the neutral theory. *J. Mol. Evol.* **21** (1984), 90-92.
912. JUNGCK, J. R., AND DICK, A. G. Computer-assisted sequencing, interval graphs, and molecular evolution. *BioSystems* **15** (1981), 259-273.
913. JUNGCK, J. R., AND FRIEDMAN, R. M. Mathematical tools for molecular genetics data: An annotated bibliography. *Bull. Math. Biol.* **46** (1984), 699-744.
914. JURETIC, N., JAUSSI, R., MATTES, U., AND CHRISTEN, P. Genes of nuclear encoded mitochondrial proteins: evidence for a variant oh the 3' splice-site consensus sequence. *Nucl. Acids Res.* **15** (1987), 10083-10086.
915. JURETIC, N., AND THEUS, M. Evidence for the existence of splicing and processing factors specific for the genes of nuclear encoded mitochondrial proteins. *Experientia* **46** (1990), A31-A31.
916. KABAT, E. A. Immunoglobulin sequence data. The generation of antibody diversity and its genetic control, (in) *The Role of Data in Scientific Progress*, Glaeser, P.S., (ed.), (Elsevier, North Holland), pp.97-102, 1985.
917. KABAT, E. A., AND WU, T. T. The influence of nearest-neighbor amino acids on the conformation of the middle amino acid in proteins: comparison of predicted and experimental determination of beta-sheets in concanavalin A. *Proc. Natl. Acad. Sci. USA* **70** (1973), 1473-1477.
918. KABAT, E. A., AND WU, T. T. The influence of nearest-neighboring amino acids residues on aspects of secondary structure of proteins. Attempts to locate alpha-helices and beta-sheets. *Biopolymers* **12** (1973), 751-754.
919. KABAT, E. A., AND WU, T. T. Further comparison of predicted and experimentally determined structure of adenylate kinase. *Proc. Natl. Acad. Sci. USA* **71** (1974), 4217-4220.
920. KABSCH, W., AND SANDER, C. Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* **22** (1983), 2577-2637.
921. KABSCH, W., AND SANDER, C. How good are predictions of protein secondary structure? *FEBS Lett.* **155** (1983), 179-182.
922. KABSCH, W., AND SANDER, C. On the use of sequence homologies to predict protein structure: identical pentapeptides can have completely different conformations. *Proc. Natl. Acad. Sci. USA* **81** (1984), 1075-1078.
923. KABSCH, W., SANDER, C., AND TRIFONOV, E. The ten helical twist angles of B-DNA. *Nucl. Acids Res.* **10** (1982), 1097-1104.
924. KAHN, P., AND CAMERON, G. EMBL data libray, (in) *Methods in Enzymology-Molecular Evolution: Computer Analysis of Protein and Nucleic Acid Sequences*, Vol.183, Doolittle, R.F., (ed.), (Academic Press, NY), pp.23-31, 1990.

925. KALAMBET, Y. A., AND ALEXANDROV, A. A. Nucleic acids functional signal database project, (in) **Abstracts of: Application of Data Bases and Microcomputers in Molecular Biology**, p.22, 1988.
926. KALAMBET, Yu. A., AND ALEXANDROV, A. A. Project of the functional signals database, (in) **Theoretical Research and Data Banks in Molecular Biology and Genetics**, (Abstracts of the 3rd National Conference), p.21, 1988.
927. KAM-MORGAN, L., GILL, B., AND MUTHUKRISHNAN, S. DNA restriction fragment length polymorphism: a strategy for genetic mapping of D genome of wheat. *Genome* **32** (1989), 724-732.
928. KAMINIR, L. B. Computers effective instruments in molecular-biological research. *Molekuljarnaya Biologiya* **13** (1979), 983-993.
929. KANAOKA, M., KISHIMOTO, F., UEKI, Y., AND UMEYAMA, H. Alignment of protein sequences using the hydrophobic core scores. *Protein Engineering* **2** (1989), 347-351.
930. KANEHISA, M. Los Alamos sequence analysis package for nucleic acids and proteins. *Nucl. Acids Res.* **10(1)** (1982), 183-196.
931. KANEHISA, M. Use of statistical criteria for screening potential homologies in nucleic acid sequences. *Nucl. Acids Res.* **12(1)** (1984), 203-213.
932. KANEHISA, M. A multivariate analysis method for discriminating protein secondary structural segments. *Protein Engineering* **2(2)** (1988), 87-92.
933. KANEHISA, M. Databases: What's there and what's needed, (in) **Biomolecular Data: A Resource in Transition**, Swartz D. G., MacDonell M. T., Colwell R. R. (eds), (Oxford University Press, NY), pp.227-231, 1989.
934. KANEHISA, M., AND DELISI, C. Prediction of protein and nucleic acid structure: Problems and prospects. *Acta Appl. Math.* **4** (1985), 114-137.
935. KANEHISA, M., FICKETT, J. W., AND GOAD, W. A relational database system for the maintenance and verification of the Los Alamos sequence library. *Nucl. Acids Res.* **12(1)** (1984), 149-158.
936. KANEHISA, M., KLEIN, P., GREIF, P., AND DELISI, C. Computer analysis and structure prediction of nucleic acids and proteins. *Nucl. Acids Res.* **12** (1984), 417-428.
937. KANEHISA, M. I. Design and application of DNA and protein data banks. *Protein Nucleic Acid and Enzyme* **29** (1986), 187-194.
938. KANEHISA, M. I., AND GOAD, W. B. Pattern recognition in nucleic acid sequences. II. An efficient method for finding locally stable secondary structures. *Nucl. Acids Res.* **10(1)** (1982), 265-278.
939. KAPITONOV, V. V. Bank of viruses, phages and plasmids genomes, (in) **Abstracts of: Application of Data Bases and Microcomputers in Molecular Biology**, (Jena), p.27, 1988.
940. KAPITONOV, V. V., AND AMIKISHIYEV, V. G. Gene conversion as a factor of preventing regulatory sites from losing validity for MGE, (in) **Modelling and Computer Methods in Molecular Biology and Genetics**, N. A. Kolchanov, V. A. Ratner, (eds.), (Novosibirsk), p.161, 1990.
941. KARIMI, E., AND BARRON, S. GenEng: A natural language interface system to GenBank. Human Genome III, San Diego, California, October 21-23 (poster abstract), 1991.
942. KARLIN, S. Comparative analysis of structural relationships in DNA and protein sequences, (in) **Evolutionary Processes and Theory**, Karlin, S., Nevo, E., (Academic Press, NY), 1986.
943. KARLIN, S. Significant potential secondary structures in the Epstein-Barr virus genome. *Proc. Natl. Acad. Sci. USA* **83** (1986), 6915-6919.
944. KARLIN, S., AND ALTSCHUL, S. F. Methods for assessing the statistical significance of molecular sequence features by using general scoring schemes. *Proc. Natl. Acad. Sci. USA* **90** (1990), 2264-2268.
945. KARLIN, S., AND ALTSCHUL, S. F. Methods for assessing the statistical significance of molecular sequence features by using general scoring schemes. *Proc. Natl. Acad. Sci. USA* **87** (1990), 2264-2268.
946. KARLIN, S., BLAISDELL, B. E., AND BRENDL, V. Identification of significant sequence patterns in proteins, (in) **Methods in Enzymology-Molecular Evolution: Computer Analysis of Protein and Nucleic Acid Sequences**, Vol.183, Doolittle, R.F., (ed.), (Academic Press, NY), pp.388-402, 1990.
947. KARLIN, S., BLAISDELL, B. E., MOCARSKI, E. S., AND BRENDL, V. A method to identify distinctive charge configurations in protein sequences, with applications to human herpes virus polypeptides. *J. Mol. Biol.* **205** (1989), 165-177.
948. KARLIN, S., AND GHANDOUR, G. Comparative statistics for DNA and protein sequences: multiple sequence analysis. *Proc. Natl. Acad. Sci. USA* **82** (1985), 6186-6190.
949. KARLIN, S., AND GHANDOUR, G. Comparative statistics for DNA and protein sequences: Single sequence analysis. *Proc. Natl. Acad. Sci. USA* **82** (1985), 5800-5804.
950. KARLIN, S., AND GHANDOUR, G. Multiple alphabet amino acid sequence comparisons of the immunoglobulin chain constant domain. *Proc. Natl. Acad. Sci. USA* **82** (1985), 8597-8601.
951. KARLIN, S., AND GHANDOUR, G. The use of multiple alphabets in kappa-gene immunoglobulin DNA sequence comparisons. *EMBO J.* **4** (1985), 1217-1223.
952. KARLIN, S., GHANDOUR, G., FOULSER, D. E., AND KORN, L. J. Comparative analysis of human and bovine papillomaviruses. *Mol. Biol. Evol.* **1(4)** (1984), 357-370.
953. KARLIN, S., GHANDOUR, G., OST, F., TAVARE, S., AND KORN, L. New approaches for computer analysis of nucleic acid sequences. *Proc. Natl. Acad. Sci. USA* **80** (1983), 5660-5664.

954. KARLIN, S., MORRIS, M., GHANDOUR, G., AND LEUNG, M.-Y. Algorithms for identifying local molecular sequence features. *CABIOS* 4(1) (1988), 41-51.
955. KARLIN, S., MORRIS, M., GHANDOUR, G., AND LEUNG, M.-Y. Efficiency algorithms for molecular sequence analysis. *Proc. Natl. Acad. Sci. USA* 85 (1988), 841-845.
956. KARLIN, S., AND OST, F. Counts of long aligned word matches among random letter sequences. *Adv. Appl. Prob.* 19 (1987), 293-351.
957. KARLIN, S., OST, F., AND BLAISDELL, B. E. Patterns in DNA and amino acid sequences and their statistical significance, (in) *Mathematical Methods for DNA Sequences*, M. Waterman, (CRC Press, Boca Raton), 1989.
958. KARP, P. A process oriented model of bacterial gene regulation. *Stanford University Knowledge Systems Laboratory Report No. KSL-88-18* (1988).
959. KARP, P. Hypothesis formation as a design, Stanford University Knowledge Systems Laboratory Report, No. KSL-89-11, 1989.
960. KARP, P., AND WANG, F. Biomedical Database: Design and Implementation. (in) *Mathematical Sciences in Genomic Analysis*, in preparation, 1991.
961. KARP, P. D. Hypothesis Formation and Qualitative Reasoning in Molecular Biology, Ph.D. Thesis, Stanford University Computer Science Department, p.319, (Technical reports STAN-CS-89-1263, KSL-89-52), June 1989.
962. KARP, P. D. The ASN.1 printfile parser and path-manipulation package. *Tech. Rep.* 5, National Center for Biotechnology Information, National Library of Medicine, Bethesda, MD, 1990.
963. KARP, P. D., AND FRIEDLAND, P. E. Coordinating the Use of Qualitative and Quantitative Knowledge in Declarative Device Modeling, (in) *Artificial Intelligence, Modeling and Simulation*, Widman, L.E., Helman, D.H., Loparo, K., (John Wiley and Sons), 1989.
964. KARP, R. M., AND RABIN, M. O. Efficient randomized pattern-matching algorithms. *IBM J. Res. Devel.* 32 (1987), 249-260.
965. KARPLUS, P. A., AND SCHULZ, G. E. Prediction of chain flexibility in proteins. *Naturwissenschaften* 72 (1985), 212-213.
966. KASHYAP, R. L., AND OOMMEN, B. J. The noisy substring matching problem. *IEEE Transactions on Software Engineering* SE-9 (1983), 365-370.
967. KATSUBE, Y. Tertiary structure and molecular evolution of protein. *Protein, Nucleic Acid and Enzyme* 29 (1986), 143-150.
968. KEIL, B. Can protein data bank help to define the binding sites of proteolytic enzymes. *Protein, Nucleic Acid and Enzyme* 29 (1986), 11-21.
969. KEIL, B. Databases in molecular biology: a CODATA task group at work. *Protein Seq. Data Anal.* 1 (1987), 123-126.
970. KEIL, B. Protein database network-an indispensable element of fundamental information in biology, (in) *Computer Handling and Dissemination of Data*, Glaeser, P.S., (ed.), (Elsevier, North Holland), pp.368-372, 1987.
971. KEIL, B. Proteolysis data bank: specificity of alpha-chymotrypsin from computation of protein cleavage. *Protein Seq. Data Anal.* 1 (1987), 13-20.
972. KEIL, B. Cooperation between databases and scientific community, (in) *Methods in Enzymology-Molecular Evolution: Computer Analysis of Protein and Nucleic Acid Sequences*, Vol.183, Doolittle, R.F., (ed.), (Academic Press, NY), pp.50-60, 1990.
973. KEIM, P., HEINRIKSON, R., AND FITCH, W. An examination of the expected degree of sequence similarity that might arise in proteins that have converged to similar conformational states. *J. Mol. Biol.* 151 (1981), 179-197.
974. KEL, A. E., KISELEVA, E. V., ISCHENKO, I. V., AND KOLCHANOV, N. A. Mathematical analysis of grand translation polysomes from chironomus salivary gland cell structural organization, (in) *Modelling and Computer Methods in Molecular Biology and Genetics*, Kolchanov, N.A., Ratner, V.A., (eds.), (Abstracts of the National Conference, Novosibirsk), pp.203-204, 1990.
975. KEL, A. E., KOLCHANOV, N. A., AND SOLOVYEV, V. Molecular mechanism of promoter functioning based on non-exact overlapping repeats, (in) *Theoretical Research and Data Banks in Molecular Biology and Genetics*, (Abstracts of the National Conference, Novosibirsk), pp.87-89, 1986.
976. KEL, A. E., PONOMARENKO, M. P., ORLOV, Yu. I., MISCHENKO, T., AND KOLCHANOV, N. A. The computer modelling of translation, (in) *Computer Analysis of Structure, Function, and Evolution of Genetic Macromolecules. Intellectualization Problems*, N.A. Kolchanov, (ed.), (Novosibirsk), pp.155-167, 1989.
977. KEL, A. E., PONOMARENKO, M. P., ORLOV, Yu. I., MISCHENKO, T., AND KOLCHANOV, N. A. The computer system of functional sites' analysis in polinucleotide sequences, (in) *Computer Analysis of Structure, Function, and Evolution of Genetic Macromolecules. Itellectualization Problems*, N.A. Kolchanov, (ed.), (Novosibirsk), pp.221-242, 1989.
978. KEL, A. E., PONOMARENKO, M. P., ORLOV, Yu. I., ISCHENKO, I. V., AND KOLCHANOV, N. A. The computer system of functional sites' analysis in polinucleotide sequences, Modelling and Computer Methods (in) *Molecular Biology and Genetics*, Kolchanov, N.A., Ratner,V.A. (eds.), (Abstracts of the National Conference, Novosibirsk), pp.51-52, 1990.

979. KELLER, C., CORCORAN, M., AND ROBERTS, R. J. Computer programs for handling nucleic acid sequences. *Nucl. Acids Res.* **12**(1) (1984), 377-395.
980. KELLER, E. B., AND NOON, W. A. Intron splicing: a conserved internal signal in introns of animal pre-mRNAs. *Proc. Natl. Acad. Sci. USA* **81** (1984), 7417-7420.
981. KELLER, E. B., AND NOON, W. A. Intron splicing: a conserved internal signal in introns of Drosophila pre-mRNAs. *Nucl. Acids Res.* **13** (1985), 4971-4981.
982. KELLER, W. Determination of the number of superhelical turns in simian virus 40 DNA by gel electrophoresis. *Proc. Natl. Acad. Sci.* **72** (1975), 4876.
983. KELLY, J. M., AND MEYER, E. F. Jr. Storage and retrieval of nucleic acid sequence data. *Computers and Chemistry* **4** (1983), 107-111.
984. KELLY, L., AND HOLLADAY, L. A. Comparison of scales of amino acid side chains properties by conservation during evolution of four proteins. *Protein Engineering* **1** (1987), 137-140.
985. KEMPF, G. J. L., AND GRAY, P. M. D. Finding hydrophobic microdomains using an object-oriented database. *CABIOS* **6**(4) (1990), 357-364.
986. KIHARA, Y., AKIYAMA, K., AND ISONO, K. The physical map of the whole *E. coli* chromosome: application of a new strategy for rapid analysis and sorting of a large genomic library. *Cell* **50** (1987), 495-508.
987. KIM, H. J., AND LEE, J. G. Partial sum problem mapping into a hypercube. *Info. Proc. Letters* **36** (1990), 221-224.
988. KIMURA, M. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *J. Mol. Evol.* **16** (1980), 111-120.
989. KIMURA, M. Estimation of evolutionary distances between homologous nucleotide sequences. *Proc. Natl. Acad. Sci. USA* **78** (1981), 454-458.
990. KIMURA, M., AND OHTA, T. On the stochastic model for estimation of mutational distance between homologous proteins. *J. Mol. Evol.* **2** (1972), 87-90.
991. KINGSBURY, D. Computational biology for biotechnology: Part 2. Applications of scientific computing in biotechnology. *Trends Biotech* **7** (1989), 126-132.
992. KINGSBURY, D. T. Computational biology for biotechnology: Part 1. The role of the computational infrastructure. *Trends Biotech* **7** (1989), 82-87.
993. KIRKWOOD, T. B. L. Identification and analysis of DNA sequence homologies, (in) *Proceedings of the XIVth International Biometric Conference*, Namur, pp.157-168, 1988.
994. KISHINO, H., AND HASEGAWA, M. Converting distance to time: Application to human evolution, (in) *Methods in Enzymology-Molecular Evolution: Computer Analysis of Protein and Nucleic Acid Sequences*, Vol.183, Doolittle, R.F., (ed.), (Academic Press, NY), pp.550-570, 1990.
995. KISHINO, H., MIYATA, T., AND HASEGAWA, M. Maximum likelihood inference of protein phylogeny and the origin of chloroplasts. *J. Mol. Evol.* **31** (1990), 151-160.
996. KLAEMBT, D., AND RICHTER, O. Computer programs for studying conformations in ribonucleic acids. *J. Theor. Biol.* **58** (1976), 319-324.
997. KLAPPER, M. H. The independent distribution of amino acid near neighbor pairs into polypeptides. *Biochem. Biophys. Res. Commun.* **78** (1977), 1018-1024.
998. KLEFFE, J. Statistical analysis of macromolecular sequences. Pattern probabilities under the dinucleotide permutation model, (in) *Statistical Data Analysis and Inference*, Dodge, Y., (Elsevier Science Publishers, North-Holland), pp.557-571, 1989.
999. KLEFFE, J., AND BORODOVSKY, M. Statistical analysis of dinucleotide compositions and its statistical application to the identification of coding regions, (in) *Modelling and Computer Methods in Molecular Biology and Genetics*, (Proceedings of the International Conference, Novosibirsk), 1990. In the press.
1000. KLEFFE, J., AND LANGBECKER, U. Exact computation of pattern probabilities in random sequences generated by Markov chains. *CABIOS* **6**(4) (1990), 347-354.
1001. KLEIN, P. Prediction of protein structural class by discriminant analysis. *Biochim. Biophys. Acta* **874** (1986), 205-215.
1002. KLEIN, P., AND DELISI, C. Prediction of protein structural class from the amino acid sequence. *Biopolymers* **25** (1986), 1659-1672.
1003. KLEIN, P., JACQUEZ, J., AND DELISI, C. Prediction of protein function by discriminant analysis. *Math. Biosci.* **81** (1986), 177-189.
1004. KLEIN, P., KANEHISA, M., AND DELISI, C. The detection and classification of membrane-spanning proteins. *Biochim. Biophys. Acta* **815** (1985), 468-476.
1005. KLEIN, P., KANEHISA, M. I., AND DELISI, C. Prediction of protein function from sequence properties: discriminant analysis of a data base. *Biochim. Biophys. Acta* **787** (1984), 221-226.
1006. KLENIN, K. V., VOLODSKII, A. V. ANSHELEVICH, V. V., DYKHNE, A. M., AND FRANK-KAMENETSKII, M. D. Computer simulation of DNA supercoiling. *J. Mol. Biol.* **217** (1991), 413-419.
1007. KLOTZ, L. C., AND BLANKEN, R. L. A practical method for calculating C evolutionary trees from sequence data. *J. Theor. Biol.* **91** (1981), 261-272.
1008. KLOTZ, L. C., KOMAR, N., BLANKEN, R. L., AND MITCHELL, R. Calculation of evolutionary trees from sequence data. *Proc. Natl. Acad. Sci. USA* **76** (1979), 4516-4520.

1009. KNEALE, G. G., AND BISHOP, M. J. Nucleic acid and protein sequence databases. *CABIOS* 1 (1985), 11-17.
1010. KNEALE, G. G., AND KENNARD, O. The EMBL nucleotide sequence data library. *Biochem. Soc. Trans.* 12 (1984), 1011-1014.
1011. KNUTH, D. E. *The art of computer programming: fundamental algorithms*, Vol.1, (Addison-Wesley), 1973.
1012. KNUTH, D. E. *The art of computer programming: searching and sorting*, Vol.3, (Addison-Wesley), 1973.
1013. KNUTH, D. E. *The art of computer programming: seminumerical algorithms*, Vol.2, (Addison-Wesley), 1981.
1014. KNUTH, D. E., MORRIS, J. H., AND PRATT, V. Fast pattern matching in strings. *SIAM J. Computing* 6 (1977), 323-350.
1015. KOHARA, Y., AKIYAMA, K., AND ISONO, K. The physical map of the whole *E. coli* chromosome: application of a new strategy for rapid analysis and sorting of a large genomic library. *Cell* 50 (1987), 495-508.
1016. KOHLI, J., AND GROSJEAN, H. Use of three termination codons: compilation and analysis of the known eukaryotic and prokaryotic translation termination sequences. *Mol. Gen. Genet.* 182 (1981), 430-439.
1017. KOHN, M. C., AND GARFINKEL, D. Computer simulation of metabolism in palmitate-perfused rat heart. II. Behavior of complete model. *Ann. Biomed. Eng.* 11 (1983), 511.
1018. KOHN, M. C., AND GARFINKEL, D. Computer simulation of metabolism in palmitate-perfused rat heart. I. Palmitate oxidation. *Ann. Biomed. Eng.* 11 (1983), 361.
1019. KOILE, K., AND OVERTON, G. A qualitative model for gene expression, (in) *Proceedings of the 1989 Summer Computer Simulation*, Proceedings of the 1989 Summer Computer Computer Simulation Conference, July, 1989.
1020. KOILE, K., AND OVERTON, G. C. A Qualitative Model for Gene Expression, (in) *Proceedings of the 1989 Summer Computer Simulation Conference*, 1989.
1021. KOLASKAR, A. S. Analysis of biomacromolecular data, (in) *Society for Computer Simulation*, Glaeser, P.S., (ed.), (Elsevier, North Holland), pp.381-383, 1987.
1022. KOLASKAR, A. S., AND REDDY, B. V. B. cDNA sequence data analysis of prokaryotic systems. *J. Biosciences* 7 (1985), 49-59.
1023. KOLASKAR, A. S., AND REDDY, B. V. B. A method to locate protein coding sequences in DNA of prokaryotic systems. *Nucl. Acids Res.* 13(1) (1985), 185-194.
1024. KOLASKER, A. S. Analysis of biomacromolecular data, (in) *Computer handing and dissemination of data*, Glaeser, P.S., (Elsevier, North Holland), pp.381-383, 1987.
1025. KOLATA, G. B. Overlapping genes: more than anomalies? *Science* 196 (1977), 1187-1188.
1026. KOLCHANOV, N. A. Data banks and packages of applied programmes and genetics, (in) *Abstracts of: Application of Data Bases and Microcomputers in Molecular Biology*, (Jena), p.24, 1988.
1027. KOLCHANOV, N. A. Computer-informational research in accordance with the Soviet Human Genome project: Problems of the intelligent systems, (in) (*Proc. of the 1st Int. Conf., Tallahassee, Florida, In the presss*, 1991).
1028. KOLCHANOV, N. A., SHINDYALOV, I. N., AND LIM, H. A. Classification of modes of evolution of gene families by computer modelling, FSU-SCRI-91-61, April 1991.
1029. KOLCHANOV, N. A., SOLOVYEV, V. V., AND ZHARKIKH, A. A. Context methods for theoretical analysis of biological macromolecules (DNA, RNA and proteins), (in) *Structure and Evolution of Genomes*, Ratner, V.A. (eds.), (VINITI, Moscow), Itogi Nauki i Tekniki, Ser. Mol. Biol., Vol.21, pp.6-37, 1985.
1030. KOLCHANOV, N. A. (ed.). Computer Analysis of Structure, Function, and Evolution of Genetic Macromolecules. Intellectualization Problems, (Abstracts of the International Conference), (Novosibirsk), 1989.
1031. KOLCHANOV, N. A. , (ed.). Computer Analysis of Structure, Function, And Evolution of Genetic Macromolecules (Academy of Sciences of the USSR, Siberian Department), 1989.
1032. KOLLMAN, P. Molecular modeling. *Ann. Rev. Phys. Chem.* 38 (1987), 303-316.
1033. KONINGS, D. A. M., van Duijn L. P., VOORMA, H. O., AND HOGEVEG, P. Evidence for use of rare codons in the dnaG gene and other regulatory genes of *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* 80 (1983), 687-691.
1034. KONOPKA, A. K. Is the information content of DNA evolutionarily significant? *J. Theor. Biol.* 107 (1984), 697-704.
1035. KONOPKA, A. K. Theory of degenerate coding and informational parmaeters of protein coding genes. *Biochimie* 67 (1985), 455-468.
1036. KONOPKA, A. K. Compilation of DNA strand exchange sites for non-homologous recombination in somatic cells. *Nucl. Acids Res.* 16(5) (1988), 1739-1758.
1037. KONOPKA, A. K. Toward mapping functional domains in indiscriminately sequenced nucleic acids: a computational approach, (in) *Structure and Methods*, Vol.1:Human Genome Initiative and DNA Recombination, Sarma, R.H., Sarma, M.H., (ed.), (Adenine Press, NY), pp.113-125, 1990.
1038. KONOPKA, A. K., AND BRENDL, V. The missense errors in protein can be controlled by selective synonymous codon usage at the level of transcription. *Biochimie* 67 (1985), 469-473.

1039. KONOPKA, A. K., AND CHATTERJEE, D. Distance analysis and sequence properties of functional domains in nucleic acids and proteins. *Gene Anal. Techn.* 5 (1988), 87-93.
1040. KONOPKA, A. K., AND OWENS, J. Complexity charts can be used to map functional domains in DNA. *Appl.* 7 (1990), 35-88.
1041. KONOPKA, A. K., REITER, J., JUNG, M., ZARLING, D. A., AND JOVIN, T. M. Concordance of experimentally mapped or predicted Z-DNA sites with positions of selected alternating purine-pyrimidine tracts. *Nucl. Acids Res.* 13(5) (1985), 1683-1701.
1042. KONOPKA, A. K., AND SARAI, A. Somatic cell illegitimate recombination via a tetraplex DNA intermediate, (in) *Human Genome Initiative and DNA Recombination*, Proceedings of The Sixth Conversation in Biomolecular Stereodynamics, Vol.1, Sarma, R.H., Sarma, M.H., (eds.), (Adenine Press, NY), 1989.
1043. KONOPKA, A. K., AND SMYTHERS, G. W. DISTAN-a program which detects significant distances between short oligonucleotides. *Comput. Appl. Biosci.* 3 (1987), 193-201.
1044. KONOPKA, A. K., SMYTHERS, G. W., OWENS, J., AND MAIZEL, J. V. Jr. Distance analysis helps to establish characteristic motifs in intron sequences. *Gene Anal. Techn.* 4 (1987), 63-74.
1045. KORN, L. J., AND QUEEN, C. Analysis of biological sequences on small computers. *DNA* 3(6) (1984), 421-436.
1046. KORN, L. J., QUEEN, C. L., AND WEGMAN, M. N. Computer analysis of nucleic acid regulatory sequences. *Proc. Natl. Acad. Sci. USA* 74 (1977), 4401-4405.
1047. KOROTKOV, E. V., AND CHALEY, M. B. A novel type of mirror symmetry and a novel wide-spread family of MBI repeats, (in) *Human Genome-90*, (Abstracts of the 1st National Conference), Bayev, A.E. (ed.), (Pereslav-Zalesky, Moscow), pp.118-119, 1990.
1048. KOTANI, M. A brief survey of the task group on coordination of protein sequence data banks and of its parent organization CODATA. *Protein, Nucleic Acid and Enzyme* 29 (1986), 8-10.
1049. KOTON, P. Towards a problem solving system for molecular genetics. Tech. Rep. 338, MIT Laboratory for Computer Science, 1985.
1050. KOZAK, M. Possible role of flanking nucleotides in recognition of the aug initiation codon by eukaryotic ribosomes. *Nucl. Acids Res.* 9 (1981), 5233-5255.
1051. KOZAK, M. Compilation and analysis of sequences upstream from the translational start site in eukaryotic mRNAs. *Nucl. Acids Res.* 12 (1984), 857-872.
1052. KOZAK, M. An analysis of 5'-noncoding sequences from 699 vertebrate messenger RNAs. *Nucl. Acids Res.* 15 (1987), 8125-8148.
1053. KOZAK, M. The scanning model for translation: An update. *J. Cell Biol.* 108 (1989), 229-241.
1054. KRAWCZAK, M. Algorithms for the restriction-site mapping of DNA molecules. *Proc. Natl. Acad. Sci. USA* 85 (1988), 7298-7301.
1055. KRAWETZ, S. A. Sequence errors described in GenBank: a means to determine the accuracy of DNA sequence interpretation. *Nucl. Acids Res.* 17 (1989), 3951-3957.
1056. KRAWETZ, S. A. Sequence interpretation and errors: What is the error rate of sequences submitted to GenBank? (in) *Genome Sequencing Conference II*, (abstracts), (Hyatt Regency, Hilton Head Island, South Carolina), 1990.
1057. KRCHNAK, V., MACH, O., AND MALÝ, A. Computer prediction of potential immunogenic determinants from protein amino acid sequence. *Anal. Biochem.* 165 (1987), 200-207.
1058. KRIGBAUM, W. R., AND KOMORIYA, A. Local interactions as a structure determinant for protein molecules: II. *Biochim. Biophys. Acta* 576 (1979), 204-228.
1059. KRIGBAUM, W. R., AND KOMORIYA, A. Local interactions as a structure determinant for protein molecules: III. *Biochim. Biophys. Acta* 576 (1979), 229-246.
1060. KRISHNAN, G., KAUL, R. K., AND JAGADEESWARAN, P. DNA sequence analysis: a procedure to find homologies among many sequences. *Nucl. Acids Res.* 14 (1986), 543-550.
1061. KRISTOFFERSON, D. The BIONET electronic network. *Nature* 325 (1987), 555-556.
1062. KRUSKAL, J. B. An anthology of algorithms and concepts for sequence comparison, (in) *Time Warps, String Edits, and Macromolecules: The Theory and Practice of Sequence Comparison*, Sankoff, D., Kruskal, J.B., (eds.), (Addison-Wesley, Reading, Massachusetts), pp.265-310, 1983.
1063. KRUSKAL, J. B. An overview of sequence comparison, (in) *Time Warps, String Edits, and Macromolecules: The Theory and Practice of Sequence Comparison*, Sankoff, D., Kruskal, J.B., (eds.), (Addison-Wesley, Reading, Massachusetts), pp.1-44, 1983.
1064. KRYSTEK, S. R. Jr., DIAS, J. A., REICHERT, L. E. Jr., AND ANDERSEN, T. T. Prediction of antigenic sites in follicle-stimulating hormones: difference profiles enhance antigenicity prediction methods. *Endocrinology* 117 (1985), 1125-1131.
1065. KRYSTEK, S. R. Jr., REICHERT, L. E. Jr., AND ANDERSEN, T. T. Analysis of computer-generated hydrophathy profiles for human glycoprotein and lactogenic hormones. *Endocrinology* 117 (1985), 1110-1124.
1066. KUBOTA, Y., NISHIKAWA, S., TAKAHASHI, S., AND OOI, T. Correspondence of homologies in amino acid sequence and tertiary structure of protein molecules. *Biochim. Biophys. Acta* 701 (1982), 242-252.
1067. KUBOTA, Y., TAKAHASHI, S., NISHIKAWA, K., AND OOI, T. Homology in protein sequences expressed by correlation coefficients. *J. Theor. Biol.* 91 (1981), 347-361.

1068. KUDO, M., LIDA, Y., AND SHIMBO, M. Syntactic pattern analysis of 5-splice site sequences of mRNA precursors in higher eukaryote genes. *CABIOS* 3 (1987), 319-324.
1069. KUHARA, S., MATSUO, F., FUTAMURA, S., FUJITA, A., SHINOHARA, T., TAKAGI, T., AND SAKAKI, Y. GENAS: A database system for nucleic acid sequence analysis. *Nucl. Acids Res.* 12 (1984), 89-99.
1070. KUHN, L. A., AND LEIGH, J. S. Jr. A statistical technique for predicting membrane protein structure. *Biochimica et Biophysica Acta*. 828 (1985), 351-361.
1071. KULKARNI, D., AND SIMON, H. A. The Processes of Scientific Discovery: The Strategy of Experimentation. *Cognitive Science* (1988), 139-175.
1072. KUNISAWA, T., HORIMOTO, K., AND OTSUKA, J. Accumulation pattern of amino acid substitutions in protein evolution. *J. Mol. Evol.* 24 (1987), 357-365.
1073. KUNTZ, I. D., CRIPPEN, G. M., KOLLMAN, P. A., AND KIMELMAN, D. Calculation of protein tertiary structure. *J. Mol. Biol.* 106 (1976), 983-994.
1074. KUZNETSOV, D. A., AND LIM, H. A. Visicoor - A program for protein zoologists, FSU-SCRI-90-198, December 1990.
1075. KUZNETSOV, D. A., AND LIM, H. A. A protein visualization program, FSU-SCRI-91-139, September 1991.
1076. KYPR, J., AND MRÁZEK, J. Reading frames of HIV genes. *J. Theor. Biol.* 141 (1989), 423-424.
1077. KYTE, J., AND DOOLITTLE, R. F. A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.* 157 (1982), 105-132.
1078. LABOUGUES, J. M. New mathematical model of the genetic code: definition and classification of mutations. *Agressologia* 19(2) (1978), 85-97.
1079. LABOUGUES, J. M. The origin and evolution of the genetic code, (in) 145th National Meeting of the American Association for the Advancement of Science, Houston, January 3-8, Herschman, A. (ed.), (A.A.A.S.), p.123, 1978.
1080. LABOUGUES, J. M. Concerning the frequencies of use of the different codons in the biosynthesis of proteins. *Agressologia* 21(5) (1980), 345-351.
1081. LABOUGUES, J. M., AND CULLMANN, G. The logic of the genetic code: optimization for the polymorphic substitution of amino acids, (in) Int. Congress of Math., (Warsaw, Poland), August 16-24, 1983.
1082. LABOUGUES, J. M., AND FIGUREAU, A. The origin and evolution of the genetic code. *Rev. Can. Biol. Exp.* 41(3) (1982), 209-216.
1083. LABOUGUES, J. M., AND GOURGAND, J. M. An example of evolutive selection against nonsense for theoretical doublet codes. *C. R. Acad. Sci. PARIS* 293(III) (1981), 665-670.
1084. LABOUGUES, J. M., AND GOURGAND, J. M. Simulation model of mutations and optimization against the occurence of terminators for fictive doublet genetic codes, (in) Proc. 1st Int. Conf. Applied Modelling and Simulation, (Lyon, France), Sept.7-11, Vol.V, pp.21-25, 1981.
1085. LADUNGA, I., CZAKO, F., CSABAÍ, I., AND GESZTI, T. Improving signal peptide predication accuracy by simulated neural network. *CABIOS* 7(4) (1991), 485-487.
1086. LAKE, J. A. A rate-independent technique for analysis of nucleic acid sequences: evolutionary parsimony. *Mol. Biol. Evol.* 4 (1987), 167-191.
1087. LALOVIC, D., AND VALJKOVIC, V. The global average DNA composition of coding regions may be determined by the electron-ion potential. *BioSystems* 23 (1990), 311-316.
1088. LAMBERT, M. H., AND SCHERAGA, H. A. Pattern recognition in the prediction of protein structure. III An importance-sampling minimization procedure. *J. Computational Chemistry* 10(6) (1989), 817-831.
1089. LANAVE, C., PREPARATA, G., SACCOME, C., AND SERIO, G. A new method for calculating evolutionary substitution rates. *J. Mol. Evol.* 20 (1984), 86-93.
1090. LANAVE, C., TOMMASI, S., PREPARATA, G., AND SACCOME, C. Transition and transversion rate in the evolution of animal mitochondrial DNA. *BioSystems* 19 (1986), 273-283.
1091. LANDAU, G. M., AND VISHKIN, U. Efficient string matching in the presence of errors, (in) IEEE Symposium on the Foundations of Computer Science, pp.126-136.
1092. LANDAU, G. M., AND VISHKIN, U. Efficient string matching with k mismatches. *Theoretical Computer Science* 43 (1986), 239-249.
1093. LANDAU, G. M., AND VISHKIN, U. Introducing efficient parallelism into approximate string matching and a new serial algorithm, (in) Proceedings of the ACM Symposium on Theory of Computing, pp.220-230, 1986.
1094. LANDAU, G. M., AND VISHKIN, U. Fast parallel and serial approximate string matching. *J. Algorithms* 10 (1989), 157-169.
1095. LANDAU, G. M., VISHKIN, U., AND NUSSINOV, R. An efficient string matching algorithm with k differences for nucleotide and amino acid sequences. *Nucl. Acids Res.* 14 (1986), 31-46.
1096. LANDAU, G. M., VISHKIN, U., AND NUSSINOV, R. An efficient string matching algorithm with k substitutions for nucleotide and amino acid sequences. *Journal of Theoretical Biology* 126 (1987), 483-490.
1097. LANDAU, G. M., VISHKIN, U., AND NUSSINOV, R. An efficient string matching algorithm with k substitutions for nucleotide and amino acid sequences. *J. Theor. Biol.* 126 (1987), 483-490.
1098. LANDAU, G. M., VISHKIN, U., AND NUSSINOV, R. Locating alignments with k differences for nucleotide and amino acid sequences. *CABIOS* 4(1) (1988). 19-24.

1099. LANDAU, G. M., VISHKIN, U., AND NUSSIONV, R. Fast alignment of DNA and protein sequences, (in) **Methods in Enzymology-Molecular Evolution: Computer Analysis of Protein and Nucleic Acid Sequences**, Vol.183, Doolittle, R.F., (ed.), (Academic Press, NY), pp.487-502, 1990.
1100. LANDER, E., MESIROV, J. P., AND TAYLOR, W. Study of protein sequence comparison metrics on the connection machine CM-2, (in) **Proceedings of the Supercomputing 88**, Vol. II, Science and Applications, Kissimmee, FL, November 14-18, 1988.
1101. LANDER, E. S. Analysis with restriction enzymes, (in) **Mathematical Methods for DNA Sequences**, M. Waterman, (ed.), (CRC Press, Boca Raton), 1989.
1102. LANDER, E. S., AND BOTSTEIN, D. Mapping complex genetic traits in humans: New methods using a complete RFLP linkage map. *Cold Spring Harbor Symposia on Quantitative Biology* LI (1986), 49-62.
1103. LANDER, E. S., AND BOTSTEIN, D. Strategies for studying heterogeneous genetic traits in humans by using a linkage map of restriction fragment length polymorphisms. *Proc. Natl. Acad. Sci. USA* 83 (1986), 7353-7357.
1104. LANDER, E. S., AND BOTSTEIN, D. Homozygosity mapping A way to map human recessive traits with the DNA of inbred children. *Science* 236 (1987), 1567-1570.
1105. LANDER, E. S., AND BOTSTEIN, D. Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics* 121 (1989), 185-199.
1106. LANDER, E. S., AND WATERMAN, M. S. Genomic mapping by fingerprinting random clones: A mathematical analysis. *Genomics* 2 (1988), 231-239.
1107. LANDRAUD, A. M., AVRIL, J. F., AND CHRETIENNE, P. An algorithm for finding a common structure shared by a family of strings. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 11 (1989), 890-895.
1108. LANGBECKER, U., KLEFFE, J., AND BORODOVSKY, M. Calculation of characteristics of asymptotic distribution of occurrences of words in genetic linguistics by Markov chain technique, (in) **Modelling and Computer Methods in Molecular Biology and Genetics**, Kolchanov, N.A., Ratner, V.A., (eds.), (Abstracts of the International Conference, Novosibirsk), p.5, 1990.
1109. LANGE, K., AND BOEHNKE, M. Some combinatorial problems of DNA restriction fragment length polymorphisms. *Am. J. Hum. Genet.* 35 (1983), 177-192.
1110. LANGLEY, P., SIMON, H. A., BRADSHAW, G. L., AND ZYTKOW, J. M. **Scientific Discovery: Computational Explorations Of The Creative Process**, (MIT Press), 1987.
1111. LAPEDES, A. How neural nets work, (in) **Neural Information Processing Systems**, Dana Z. Anderson, (ed.), (American Institute of Physics, NY), 1988.
1112. LAPEDES, A., BARNES, C., BURKES, C., FARBER, R., AND SIROTKIN, K. Application of neural networks and other machine learning algorithms to DNA sequence analysis, (in) **Computers and DNA**, SFI Studies in the Sciences of Complexity, Vol. VII, G. Bell, T. Marr, (eds.), (Addison-Wesley, MA), 1989.
1113. LAPEDES, A., BARNES, C., BURKS, C., FARBER, R., AND SIROTKIN, K. Application of neural networks and other machine learning algorithms to DNA sequence analysis, (in) **Computers and DNA**, Bell G.I., Marr T., (eds.), (Addison-Wesley, Reading), PP.157-182, 1989.
1114. LATHE, R. Synthetic oligonucleotide probes deduced from amino acid sequence data. Theoretical and practical considerations. *J. Mol. Biol.* 183 (1985), 1-12.
1115. LATHE, R., AND FINDLAY, R. Machine readable DNA sequences. *Nature* 331 (1984), 610-612.
1116. LATHE, R., AND FINDLAY, R. Reply to hamori's letter. *Nature* 314 (1985), 585-586.
1117. LATHROP, R. H., WEBSTER, T. A., AND SMITH, T. F. Ariadne: Pattern-directed inference and hierarchical abstraction in protein structure recognition. *CACM* 30 (1987), 909-921.
1118. LATHROP, R. H., WEBSTER, T. A., SMITH, T. F., AND WINSTON, P. H. ARIEL: A massively parallel symbolic learning assistant for protein structure/function, preprint, 1989.
1119. LAWRENCE, C. B. Data structures for DNA sequence manipulation. *Nucl. Acids Res.* 14 (1986), 205-216.
1120. LAWRENCE, C. B. Use of homology domains in sequence similarity detection, (in) **Methods in Enzymology-Molecular Evolution: Computer Analysis of Protein and Nucleic Acid Sequences**, Vol.183, Doolittle, R.F., (ed.), (Academic Press, NY), pp.133-146, 1990.
1121. LAWRENCE, C. B., AND GOLDMAN, D. A. Definition and identification of homology domains. *CABIOS* 4(1) (1988), 25-33.
1122. LAWRENCE, C. B., GOLDMAN, D. A., AND HOOD, R. T. Optimized homology searches of the gene and protein sequence data banks. *Bull. Math. Biol.* 48 (1986), 569-583.
1123. LAWRENCE, C. E., AND BRYANT, S. H. Hydrophobic potentials from statistical analysis of protein structures, (in) **Methods in Enzymology-Molecular design and Modeling: Concepts and Applications**, Part A, Vol.202, Langone, J.J., (ed.), (Academic Press, NY), pp.20-31, 1991.
1124. LAWRENCE, C. E., AND REILLY, A. A. Maximum likelihood estimation of subsequence conservation. *J. Theor. Biol.* 113 (1985), 425-439.
1125. LAWRENCE, C. E., AND REILLY, A. A. An expectation maximization (EM) algorithm for the identification and characterization of common sites in unaligned biopolymer sequences. *Proteins* 7 (1990), 41-51.
1126. LAWTON, J. R., MARTINEZ, F. A., AND BURKS, C. Overview of the LiMB database. *Nucl. Acids Res.* 17(15) (1989), 5885-5899.
1127. LAZOU, C. **Supercomputers and their Use**, (Oxford University Press, Cambridge), 1988.

1128. LE, S.-Y., CHEN, J.-H., AND MAIZEL, J. V. Jr. Detection of unusual RNA folding regions in HIV and SIV sequences. *CABIOS*, in press.
1129. LE, S.-Y., CHEN, J.-H., NUSSINOV, R., AND MAIZEL, J. V. Jr. An improved secondary structure computation method and its application to intervening sequences in the human alpha-like globin mRNA precursors. *CABIOS* 4(3) (1988), 337-344.
1130. LE, S.-Y., CURREY, K. M., NUSSINOV, R., AND MAIZEL, J. V. Jr. Studies of frequently recurring substructures in human alpha-like globin mRNA precursors. *Comput. Methods. Med.* 20 (1987), 563-582.
1131. LE, S.-Y., HSIANG, J. H., BRAUN, M. J., GONDA, M. A., AND MAIZEL, J. V. Stability of RNA stem-loop structure and distribution of non-random structure in the human immunodeficiency virus (HIV-I). *Nucl. Acids Res.* 16(11) (1988), 5153-5168.
1132. LE, S.-Y., AND JIANG, S.-P. The folding and dynamic process of RNA. *Acta Biophysica Sinica* 3 (1987), 227-231.
1133. LE, S.-Y., AND MAIZEL, J. V. A method for assessing the statistical significance of RNA folding. *J. Theor. Biol.* 138 (1989), 495-510.
1134. LE, S.-Y., MALIM, M. H., CULLEN, B. R., AND MAIZEL, J. V. A highly conserved RNA folding region coincident with the Rev response element of primate immunodeficiency viruses. *Nucl. Acids Res.* 18 (1990), 1613-1623.
1135. LE, S.-Y., NUSSINOV, R., AND MAIZEL, J. V. Tree graphs of RNA secondary structures and their comparisons. *Comput. Biomed. Res.*, 1989. in press.
1136. LE, S.-Y., OWENS, J., NUSSINOV, R., CHEN, J.-H., SHAPIRO, B., AND MAIZEL, J. V. Jr. RNA secondary structures: Comparison and determination of frequently recurring substructures by consensus. *CABIOS* 5 (1989), 205-210.
1137. LEE, B., AND RICHARDS, F. M. The interpretation of protein structures: estimation of static accessibility. *J. Mol. Biol.* 55 (1971), 379-400.
1138. LEMESLE-VARLOOT, L., HENRISSAT, B., GABORIAUD, C., BISSEERY, V., MORGAT, A., AND MORNON, J. P. Hydrophobic cluster analysis: procedures to derive structural and functional information from 2-D-representation of protein sequences. *Biochimie*. 72 (1990), 555-574.
1139. LENNON, G. G., AND NUSSINOV, R. Homonyms, synonyms and mutations of the sequence/structure vocabulary. *J. Mol. Biol.* 175 (1984), 425-430.
1140. LENSTRA, J. A. Evaluation of secondary structure predictions in proteins. *Biochim. Biophys. Acta* 491 (1977), 333-338.
1141. LERMAN, L. S., AND SILVERSTEIN, K. Computational simulation of DNA melting and its application to denaturing gradient gel electrophoresis, (in) *Methods in Enzymology*, Vol.155, (Academic Press, NY), pp.482-501, 1987.
1142. LESK, A. M. The analysis of protein structures: new insights from a growing data base. *BioEssays* 1 (1984), 105-110.
1143. LESK, A. M. Coordination of sequence data. *Nature* 314 (1985), 318-319.
1144. LESK, A. M. Protein structure, evolution, and conformational change. *Protein, Nucleic Acid and Enzyme* 29 (1986), 123-135.
1145. LESK, A. M. Three-dimensional structures of proteins: archives, research and applications. *Protein, Nucleic Acid and Enzyme* 29 (1986), 101-110.
1146. LESK, A. M. The Biocomputing program at EMBL. *Trends Biotechnology* 5 (1987), 317-318.
1147. LESK, A. M. Conclusion: what shall one do with a newly-determined sequence? (in) *Computational Molecular Biology. Sources and Methods for Sequence Analysis*, Lesk, A.M., (ed.), (Oxford University Press, Oxford), pp.216-219, 1988.
1148. LESK, A. M., BOSWELL, D. R., LESK, V. I., LESK, V. E., AND BAIROCH, A. A cross-reference table between the protein data bank of macromolecular structures and the National Biomedical Research Foundation Protein Identification Resource amino acid sequence data bank. *Protein Seq. Data Anal.* 2 (1989), 295-308.
1149. LESK, A. M., AND HARMAN, K. D. Computer-generated pictures of proteins, (in) *Methods in Enzymology*, Vol.115, (Academic Press, NY), pp.381-390, 1985.
1150. LESK, A. M., LEVITT, M., AND CHOHTHIA, C. Alignment of the amino acid sequences of distantly related proteins using variable gap penalties. *Protein Engineering* 1 (1986), 77-78.
1151. LESK, A. M., (ed.). *Computational Molecular Biology: Sources and Methods for Sequence Analysis*, (Oxford University Press, NY), 1988.
1152. LEVENE, S., AND CROTHERS, D. M. A computer graphics study of sequence-directed bending in DNA. *J. Biomol. Struct. Dyn.* 1 (1983), 429-435.
1153. LEVIN, J. M., ROBSON, B., AND GARNIER, J. An algorithm for secondary structure determination in proteins based on sequence similarity. *FEBS Lett.* 205 (1986), 303-308.
1154. LEVINSON, B., KENWRICK, S., LAKICH, D., HAMMONDS, G. Jr., AND GITSCHEIER, J. A transcribed gene in an intron of the human factor VIII gene. *Genomics* 7 (1990), 1-11.
1155. LEVITT, M. A simplified representation of protein conformations for rapid simulation of protein folding. *J. Mol. Biol.* 104 (1976), 59-107.

1156. LEVITT, M. Conformational preferences of amino acid in globular proteins. *Biochemistry* **17** (1978), 4277-4285.
1157. LEVITT, M. Computer simulation of DNA double-helix dynamics. *Cold Spring Harbor Symp. Quant. Biol.* **47** (1983), 251-262.
1158. LEVITT, M. Molecular dynamics of native proteins: I. Computer simulation of trajectories. *J. Mol. Biol.* **168** (1983), 595-620.
1159. LEVITT, M. Protein folding by restrained energy minimization and molecular dynamics. *J. Mol. Biol.* **170** (1983), 723-764.
1160. LEVITT, M., AND CHOTHIA, C. Structural patterns in globular proteins. *Nature* **261** (1976), 552-558.
1161. LEVITT, M., AND WARSHEL, A. Computer simulation of protein folding. *Nature* **253** (1975), 694-698.
1162. LEWIN, R. National networks for molecular biologists. *Science* **223** (1984), 1379-1380.
1163. LEWONTIN, R. C. Inferring the number of evolutionary events from DNA coding sequence differences. *Mol. Biol. Evol.* **6** (1989), 15-32.
1164. LI, H., LI, Y., AND ZHAO, H. Fractal analysis of protein chain conformation. *Int. J. Biol. Macromol.* **12(1)** (1990), 6-8.
1165. LI, M. Towards a DNA sequencing theory. Tech. Rep. CS-90-14, University of Waterloo, 1990.
1166. LI, M. Towards a DNA sequencing theory, (in) *IEEE Symposium on the Foundations of Computer Science*, pp.125-134, 1990.
1167. LI, W.-H. Models of nearly neutral mutations with particular implications for nonrandom usage of synonymous codons. *J. Mol. Evol.* **24** (1987), 337-345.
1168. LI, W.-H. A statistical test of phylogenies estimated from sequence data. *Mol. Biol. Evol.* **6(4)** (1989), 424-435.
1169. LI, W.-H., AND GOUY, M. Statistical tests of molecular phylogenies, (in) *Methods in Enzymology-Molecular Evolution: Computer Analysis of Protein and Nucleic Acid Sequences*, Vol.183, Doolittle, R.F., (ed.), (Academic Press, NY), pp.645-659, 1990.
1170. LI, W.-H., WU, C. I., AND LUO, C. C. A new method for estimating synonymous and nonsynonymous rates of nucleotide substitution considering the relative likelihood of nucleotide and codon changes. *Mol. Biol. Evol.* **2** (1985), 150-174.
1171. LIDA, Y. DNA sequences and multivariate statistical analysis. Categorical discrimination approach to 5 splice site signals of mRNA precursors in higher eukaryotes' genes. *CABIOS* **3** (1987), 93-98.
1172. LIEBIG, H. D., AND RUGER, W. Bacteriophage T4 early promoter regions. Consensus sequences of promoters and ribosome-binding sites. *J. Mol. Biol.* **208** (1989), 517-536.
1173. LIEBMAN, M. N. Molecular modeling of protein structure and function: a bioinformatic approach. *J. Comp.-Aided Molec. Design* **1** (1987), 323-341.
1174. LIEBMAN, M. N. Analysis of the biomacromolecular architecture of eukaryotic and prokaryotic serine proteases. *Journal of Industrial Microbiology* **3** (1988), 127-137.
1175. LIEBMAN, M. N., AND BRUGGE, A. L. GPGA: A prototype for analysis of the sequence-structure-function relationship in the human genome project, (in) *The Interface Between Computational Science and Nucleic Acid Sequencing*, in press.
1176. LIHOSHVAI, V. A. Rare codons: fortuity or regularity? (in) *Modelling and Computer Methods in Molecular Biology and Genetics*, Kolchanov, N.A., Ratner, V.A., (eds.), (Abstracts of the International Conference, Novosibirsk), pp.202-203, 1990.
1177. LIM, H. A. Classification of functional regions of DNA sequences using fractal representation, FSU-SCRI-91C-90, July 1991.
1178. LIM, H. A., AND Janse Van Rensburg, E. J. Electrophoresis of circular & knotted polymers/DNA, FSU-SCRI-91C-89, July 1991.
1179. LIM, H. A., AND Janse van Rensburg, E. J. A numerical simulation of electrophoresis of knotted DNA, FSU-SCRI-91-163, October 1991.
1180. LIM, H. A., W., S. G., AND JAAN, N. A model of the DNA transient orientation overshoot during gel electrophoresis. *J. Chem. Phys.* **92(1)** (1990), 709-721.
1181. LIM, V. I. Algorithms for prediction of alpha-helical and beta-structural regions in globular proteins. *J. Mol. Biol.* **88** (1974), 873-894.
1182. LIM, V. I. Polypeptide chain folding through a highly helical intermediate as a general principle of globular protein structure formation. *FEBS Lett.* **89** (1978), 10-14.
1183. LINDSAY, R., BUCHANAN, B. G., FEIGENBAUM, E. A., AND LEDERBERG, J. *Applications of Artificial Intelligence for Organic Chemistry: The DENDRAL Project*, (McGraw-Hill, NY), 1980.
1184. LINSLEY, J., AND YOUNG, S. W. Mobility models and experimental data for lambda phage concatamers during field inversion gel electrophoresis, (in) *First International Conference on Electrophoresis, Supercomputing, and The Human Genome*, Cantor, C.R. and Lim, H.A., (eds.), pp.123-156, 1990.
1185. LIPMAN, D., AND PEARSON, W. R. Rapid and sensitive protein similarity searches. *Science* **227** (1985), 1435-1441.
1186. LIPMAN, D. J., ALTSCHUL, S. F., AND KECECIYOLU, J. D. A tool for multiple sequence alignment. *Proc. Natl. Acad. Sci. USA* **86** (1989), 4412-4415.

1187. LIPMAN, D. J., AND MAIZEL, J. Comparative analysis of nucleotide acid sequences by their general constraints. *Nucl. Acids Res.* **10** (1982), 2723-2739.
1188. LIPMAN, D. J., SMITH, T. F., BECKMAN, R. J., AND WATERMAN, M. S. Hierarchical analysis of influenza: A hemagglutinin gene sequence. *Nucl. Acids Res.* **10** (1982), 5375-5389.
1189. LIPMAN, D. J., AND WILBUR, W. J. Contextual constraints on synonymous codon choice. *J. Mol. Biol.* **163** (1983), 363-376.
1190. LIPMAN, D. J., AND WILBUR, W. J. Interaction of silent and replacement changes in eukaryotic coding sequences. *J. Mol. Evol.* **21** (1985), 161-167.
1191. LIPMAN, D. J., AND WILBUR, W. J. Modelling neutral and selective evolution of protein folding. *Proc. R. Soc. Lond. B.* **245** (1991), 7-11.
1192. LIPMAN, D. J., WILBUR, W. J., SMITH, T. F., AND WATERMAN, M. S. On the statistical significance of nucleic acid similarities. *Nucl. Acids Res.* **12(1)** (1984), 215-226.
1193. LIPTON, R., MARR, T. G., AND WELSH, J. D. Computational approaches to discovering semantics in molecular biology. *Proc. of the IEEE* **77(7)** (1989), 1056-1060.
1194. LIPTON, R. J., AND LOPRESTI, D. A systolic array for rapid string comparison, (in) *Chapel Hill Conference on Very Large Scale Integration*, H. Fuchs, (ed.), (Computer Science Press, Rockville, Maryland), 1985.
1195. LIQUORI, A. M., RIPAMONTI, A., SADUN, C., OTTANI, S., AND BRAGA, D. Pattern recognition of sequence similarities in globular proteins by Fourier analysis: a novel approach to molecular evolution. *J. Mol. Evol.* **23** (1986), 80-87.
1196. LITTLE, J. W., AND MOUNT, D. W. Creating new restriction sites by silent changes in coding sequences. *Gene* **32** (1984), 67-73.
1197. LITWIN, S. The lengths of fragments covering a fixed marker in randomly cut linear and circular DNA. *J. Appl. Prob.* **16** (1979), 721-731.
1198. LIVINGSTONE, D. J. Pattern Recognition methods in rational drug design, (in) *Methods in Enzymology-Molecular Design and Modeling: Concepts and Applications*, Part B, Vol.203, Langone, J.J., (ed.), (Academic Press, NY), pp.613-638, 1991.
1199. LOCKER, J., AND BUZZARD, G. A dictionary of transcription control sequences, *DNA Sequence*, in press.
1200. LOMBARDI, S., SEIDELL, H., PULFORD, S., DUTTON, W., AND PAREKH, S. Computer programs in nucleic acid synthesis: synthetic strategy development using solid-phase chemical techniques with data storage, retrieval and analysis capabilities. *Nucl. Acids Res.* **12** (1984), 437-446.
1201. LONSDALE, D. M., HODGE, T. P., AND STOEHR, P. J. A computer program for the management of small cosmid banks. *Nucl. Acids Res.* **12** (1984), 429-436.
1202. LOOMIS, W. F., AND GILPIN, M. E. Multigene family and vestigial sequences. *Proc. Natl. Acad. Sci. USA* **83** (1986), 2143-2147.
1203. LOUIS, B. G., AND GANOZA, M. C. Signals determining translational start-site recognition in eukaryotes and their role in prediction of genetic reading frames. *Mol. Biol. Rep.* **13** (1988), 103-115.
1204. LU, P., CHEUNG, S., AND ARNDT, K. Possible molecular detent in the dna structure at regulatory sites. *J. Biomol. Struct. Dynam.* **1** (1983), 509-521.
1205. LUCHNIK, A. N., AND GLASER, V. M. DNA topological linking numbers in malignantly transformed Syrian hamster cells. *Mol. Gen. Genet.* **183** (1981), 553.
1206. LUCKOW, V. A., LITTLEWOOD, R. K., AND ROUND, R. H. Interactive computer programs for the graphic analysis of nucleotide sequence data. *Nucl. Acids Res.* **12(1)** (1984), 665-673.
1207. LUKASHIN, A. V., ANSHELEVICH, V. V., AMIRIKYAN, B. R., GRAGEROV, A. I., AND FRANK-KAMENETSKII, M. D. Neural network models for promoter recognition. *J. Biomol. Struct. Dyn.* **6** (1989), 1123-1133.
1208. LUKINA, N. I., AND SOIDL, T. R. Large stereospecific anomalies of the DNA base composition as an indication for DNA-protein complexes, (in) *Modelling and Computer Methods in Molecular Biology and Genetics*, Kolchanov, N.A., Ratner, V.A., (eds.), (Abstracts of the International Conference, Novosibirsk), pp.38-39, 1990.
1209. LUMSDEN, C. J. Gene-culture coevolution: a test of the steady-state hypothesis for gene-culture translation. *J. Theor. Biol.* **130** (1988), 391-406. (abstract only).
1210. LUO, L., AND LI, H. The statistical correlation of nucleotides in protein-coding DNA sequences. *Bull. Math. Biol.* **53(3)** (1991), 345-353.
1211. LUO, L. F. The distribution of nucleic acids in the genetic code. *Origins of Life* **19** (1989), 621-632.
1212. LUO, L. F., TSAI, L., AND ZHOU, Y. M. Informational parameters of nucleic acid and molecular evolution. *J. Theor. Biol.* **130** (1988), 351-361.
1213. LUTCKE, H. A., CHOW, K. C., MICHEL, F. S., MOSS, K. A., KERN, H. F., AND SCHEELE, G. A. Selection of AUG initiation codons differs in plants and animals. *EMBO J.* **6** (1987), 43-48.
1214. LYALL, A., HIDD., HILL, C., COLLINS, J. F., AND COULSON, A. F. W. Implementation of inexact string matching algorithms on the ICL DAP, (in) *Parallel Computing 85*, Feilmeyer, M., Joubert, G., Schendel, U., (eds.), (North-Holland, Amsterdam), pp.235-240, 1986.
1215. MACCHIATO, M., AND TRAMONTANO, A. Markov analysis of alpha-helical, beta-sheet and random coil regions of proteins. *Lettore Al Nuovo Cimento* **37(9)** (1983), 89-94.

1216. MACCHIATO, M. F., CUOMO, V., AND TRAMONTANO, A. Determination of the autocorrelation orders of proteins. *Eur. J. Biochem.* **149** (1985), 375-379.
1217. MAES, M. On a cyclic string-to-string correction problem. *Info. Proc. Letters* (1990), 73-78.
1218. MAICAS, E., AND FRIERSEN, J. D. A sequence pattern that occurs at transcription initiation region of yeast polymerase II promoters. *Nucl. Acids Res.* **18** (1990), 3387-3393.
1219. MAIER, D. The complexity of some problems on subsequences and supersequences. *J. ACM* **25** (1978), 322-336.
1220. MAIN, M. G., AND LORENTZ, R. J. An $\theta(n \log n)$ algorithm for finding all repetitions in a string. Tech. Rep. CU-CS-241-82, University of Colorado, 1982.
1221. MAIZEL, J. V. Jr. Stability of RNA stem-loop structure and distribution of non-random structure in the human immunodeficiency virus (HIV-I). *Nucl. Acids Res.* **16(11)** (1988), 5153-5168.
1222. MAIZEL, J. V. Jr. Supercomputing in Biomedical Research. *Cray Channels* **10** (1988), 3.
1223. MAIZEL, J. V. Jr., AND LENK, R. P. Enhanced graphic matrix analysis of nucleic acid and protein sequences. *Proc. Natl. Acad. Sci.* **78** (1981), 7665-7669.
1224. MAJOR, F., FELDMANN, R., AND LAPALME, G. FUS: a system to simulate conformational changes in biological macromolecules. *CABIOS* **4** (1988), 445-451.
1225. MALOY, S., AND OLSON, S. Educational software for the analysis of DNA and protein sequences. *Academic Computing* **3** (1989), 18.
1226. MALTHIERY, B., BELLON, B., GIORGI, D., AND JACQ, B. Apple II Pascal programs for molecular biologists. *Nucl. Acids Res.* **12** (1984), 569-579.
1227. MANAVALAN, P., AND PONNUSWAMY, P. K. Hydrophobic character of amino acid residues in globular proteins. *Nature* **275** (1978), 673-674.
1228. MANBER, U., AND BAEZA-YATES, R. An algorithm for string matching with a sequence of don't cares. *Info. Proc. Letters* **37** (1991), 133-136.
1229. MANBER, U., AND MYERS, E. Suffix arrays: a new method for on-line string searches. *siam journal on computing*, in press, 1990.
1230. MANDLER, J. ANTIGEN: protein surface residue prediction. *CABIOS* **4** (1988), 493.
1231. MANDLER, J. HYSTRUC: hydropathy and secondary structure prediction. *CABIOS* **4** (1988), 309.
1232. MANNING, G. The use of the DAP, a massively parallel computing system, for information retrieval and processing. IEEE Colloquium on Parallel Techniques for Information Retrieval Report, UK18.13, 1989.
1233. Mapping The Human Genome. Lawrence Berkeley Laboratory, University of California, 1989.
1234. MARCAUD, H., GABARRO-ARPA, J., EHRLICH, R., AND REISS, C. An algorithm for studying cooperative transitions in DNA. *Nucl. Acids Res.* **14** (1986), 551-558.
1235. MARCK, C. Fast analysis of DNA and protein sequence on Apple IIe: restriction sites search, alignment of short sequence and dot matrix analysis. *Nucl. Acids Res.* **14** (1986), 583-590.
1236. MARGALIT, H., SHAPIRO, B. A., NUSSINOV, R., OWENS, J., AND JERNIGAN, R. L. Helix stability in prokaryotic promoter regions. *Biochemistry* **27** (1988), 5179-5188.
1237. MARGALIT, H., SPOUGE, J. L., CORNETTE, J. L., CEASE, K. B., DELISI, C., AND BERZOFSKY, J. A. Prediction of immunodominant helper T cell antigenic sites from the primary sequence. *J. Immunol.* **138** (1987), 2213-2229.
1238. MARIN, A., BERTRANPETIT, J., OLIVER, J. L., AND MEDINA, J. R. Variation in G+C-content and codon choice: differences among synonymous codon groups in vertebrate genes. *Nucl. Acids Res.* **17** (1990), 6181-6189.
1239. MARKOWITZ, E. Estimation and testing goodness-of-fit for some models of codon fixation variability. *Biochem. Genet.* **4** (1970), 595-601.
1240. MARLIERE, P. Computer building and folding of fictitious transfer-RNA sequences. *Biochimie* **65** (1983), 267-273.
1241. MAROUN, L. E., AND ADAMS, R. B. Evolutionary conservation and multiple presence of dual-site 18S rRNA-mRNA complimentarity in eukaryotic viruses, (in) *Proc. 1987 Miami Winter Symp.*, (Cambridge University Press), 1987.
1242. MAROUN, L. E., DEGNAR, M., PRECUP, J. W., AND FRANCISKOVICH, P. P. Eukaryotic mRNA 5'-leader sequences have dual regions of complementarity to the 3'-terminus of 18s RNA. *J. Theor. Biol.* **120** (1986), 85-98.
1243. MARTINDALE, D. W. Codon usage in Tetrahymena and other ciliates. *J. Protozool.* **36** (1989), 29-34.
1244. MARTINEZ, H. M. An efficient method for finding repeats in molecular sequences. *Nucl. Acids Res.* **11(13)** (1983), 4629-4634.
1245. MARTINEZ, H. M. An RNA folding rule. *Nucl. Acids Res.* **12(1)** (1984), 323-334.
1246. MARTINEZ, H. M. A flexible multiple sequence alignment program. *Nucl. Acids Res.* **16** (1988), 1683-1691.
1247. MARTINEZ, H. M. An RNA secondary structure workbench. *Nucl. Acids Res.* **16** (1988), 1789-1798.
1248. MARTINEZ, H. M. Detecting pseudoknots and other local base-pairing structures in RNA sequences, (in) *Methods in Enzymology-Molecular Evolution: Computer Analysis of Protein and Nucleic Acid Sequences*, Vol.183, Doolittle, R.F., (ed.), (Academic Press, NY), pp.306-317, 1990.

1249. MARTINEZ, H. M., (ed.). Mathematical and computational problems in the analysis of molecular sequences. *Bull. Math. Biol.* 46 (1984), 461-744.
1250. MARUYAMA, T., GOJOBORI, T., AOTA, S., AND IKEMURA, T. Codon usage tabulated from the genbank genetic sequence data. *Nucl. Acids Res.* 14 (1986), r151-r197.
1251. MARVEL, C. C. A program for the identification of tRNA-like structures in DNA sequence data. *Nucl. Acids Res.* 14 (1986), 431-435.
1252. MATASSI, G., MONTERO, L. M., SALINAS, J., AND BERNARDI, G. The isochore organization and the compositional distribution of homologous coding sequences in the nuclear genome of plants. *Nucl. Acids Res.* 17 (1989), 5273-5290.
1253. MATOUSCHEK, A., AND FERSHT, A. R. Protein engineering in analysis of protein folding pathways and stability, (in) *Methods in Enzymology-Molecular design and Modeling: Concepts and Applications*, Part A, Vol.202, Langone, J.J., (ed.), (Academic Press, NY), pp.82-112, 1991.
1254. MATTERN, M. R., AND PAINTER, R. B. Dependence of mammalian DNA replication upon DNA super-coiling. *Biochim. Biophys. Acta* 583 (1979), 293.
1255. MATTHEWS, B. W. Comparison of the predicted and observed secondary structure of T4 phage lysozyme. *Biochim. Biophys. Acta* 405 (1975), 442-451.
1256. MATTHEWS, B. W., AND G., R. M. Comparisons of protein structures, (in) *Methods in Enzymology*, Vol.115, (Academic Press, NY), pp.397-420, 1985.
1257. MATTHEWS, B. W., AND GROSS, A. J. Modeling DNA: Some applications of error correcting codes and information theory. ABS #80T-C35, 1980.
1258. MAULIK, S. Protein databases and software on BIONET. *Protein Seq. Data Anal.* 2 (1989), 111-114.
1259. MAURER, S. B. A minimum cycle problem in DNA research. *Notices Amer. Math. Soc.* 69 (1979), A-448.
1260. MAVROVOUNIOTIS, M. Group contributions for estimating standard Gibbs energies of formation of biochemical compounds in aqueous solution. *Biotechnology and Bioengineering* 36 (1990), 1070-1082.
1261. MAVROVOUNIOTIS, M. L. Computer-Aided Design of Biochemical Pathways, Ph.D. Thesis, Massachusetts Institute of Technology, 1989.
1262. MAXAM, A. M., AND GILBERT, W. A new method for sequencing DNA. *Proc. Natl. Acad. Sci.* 74(2) (1977), 560-564.
1263. MAXFIELD, F. R., AND SCHERAGA, H. A. Status of empirical methods for the prediction of protein backbone topography. *Biochemistry* 15 (1976), 5138-5153.
1264. MAXFIELD, F. R., AND SCHERAGA, H. A. Improvements in the prediction of protein backbone topography by reduction of statistical errors. *Biochemistry* 18 (1979), 697-704.
1265. MCBRYAN, O. A. Overview of Current Developments in Parallel Architectures, (in) *Parallel Supercomputing: Methods, Algorithms and Applications*, G.F. Carey, (ed.), (John Wiley, NY).
1266. MCCALDON, P., AND ARGOS, P. Oligopeptide biases in protein sequences and their use in predicting protein coding regions in nucleotide sequences. *Proteins* 4 (1988), 99-122.
1267. MCCALDON, P., AND ARGOS, P. Oligopeptide biases in protein sequences and their use in predicting protein coding regions in nucleotide sequences. *Structure, Function, and Genetics* 4 (1989), 99-122.
1268. McCALLUM, D., AND SMITH, M. Computer processing of DNA sequence data. *J. Mol. Biol.* 116 (1977), 29-30.
1269. MCCLELLAND, M. Selection against dam methylation sites in the genomes of DNA of bacteriophages. *J. Mol. Evol.* 21 (1985), 317-322.
1270. MCCREIGHT, E. M. A space-economical suffix tree construction algorithm. *J. ACM* 23 (1976), 262-272.
1271. McDONNELL, M. T., AND COLWELL, R. R. A microcomputer program for the computation of free energy of the secondary structure of RNA oligomers by Ninio's rules. *J. Microbiological Methods* 3 (1985), 33-42.
1272. McDONNELL, M. T., AND COLWELL, R. R. Dot-matrix analysis of nucleic acid sequences: an interactive microcomputer program. *J. Microbiological Methods* 3 (1985), 247-254.
1273. MCGREGOR, M. J., FLORES, T. P., AND STERNBERG, M. J. E. Prediction of β -turns in proteins using neural networks. *Protein Engineering* 2 (1989), 521-526.
1274. MCKEAN, D. J., PETERS, E. H., WALDBY, J. I., AND SMITHIES, O. Amino acid sequence determination with radioactive proteins. *Biochemistry* 13 (1974), 3048-3051.
1275. MCKUSICK, V. A. The Human Genome Organization: History, Purposes, and Membership. *Genomics* 5 (1989), 385-387.
1276. McLACHLAN, A. D. Test for comparing related amino acid sequences. Cytochrome c and cytochrome c551. *J. Mol. Biol.* 61 (1971), 409-424.
1277. McLACHLAN, A. D. Repeating sequences and gene duplication in proteins. *J. Mol. Biol.* 64 (1972), 417-437.
1278. McLACHLAN, A. D. Analysis of periodic patterns in amino acid sequences: Collagen. *Biopolymers* 16 (1977), 1271-1297.
1279. McLACHLAN, A. D., AND BOSWELL, D. R. Confidence limits for homology in protein or gene sequences. The c-myc oncogene and a denovirus E1A protein. *J. Mol. Biol.* 185 (1985), 39-49.
1280. McLACHLAN, A. D., STADEN, R., AND BOSWELL, D. R. A method for measuring the non-random bias of a codon usage table. *Nucl. Acids Res.* 12 (1984), 9567-9575.

1281. MC LAUGHLAN, J., GAFFNEY, D., WHITTON, J. L., AND CLEMENS, J. B. The consensus sequence YGT-GTTYY located downstream from the AATAAA signal is required for efficient formation of mRNA 3' termini. *Nucl. Acids Res.* **13** (1985), 1347-1368.
1282. MC LAUGHLIN, P. J., AND DAYHOFF, M. O. Evolution of species and proteins: a time scale, (in) *Atlas of Protein Sequence and Structure*, 5:47-52, (Nat. Biomed. Res. Found, Washington D.C.), 1972.
1283. McMORRIS, F. R., AND POWERS, R. C. Consensus weak hierarchies. *Bull. Math. Biol.* **53**(5) (1991), 679-684.
1284. MCPHERSON, D. T. Codon preference reflects mistranslational constraints. *Nucl. Acids Res.* **16** (1988), 4111-4120.
1285. MECHER, U. A readable and space-efficient DNA sequence representation: Application to caulimoviral DNAs. *CABIOS* **4** (1988), 93-96.
1286. MEDNIKOV, B. The analogy. *Nauka i Zhizn* **10** (1989), 70-75.
1287. MELCHER, U. A readable and space-efficient DNA sequence representation: application to caulimoviral DNAs. *Comput. Appl. Biosci.* **4** (1988), 93-96.
1288. MENENDEZ-ARIAS, L., GOMEZ-GUTIERREZ, J., GARCIA-FERNANDEZ, M., GARCIA-TEJEDOR, A., AND MORAN, F. A BASIC microcomputer program to calculate the secondary structure of proteins from the circular dichroism spectrum. *CABIOS* **4** (1988), 479-482.
1289. MENGERITSKY, G., AND SMITH, T. F. Recognition of characteristics patterns in sets of functionally equivalent DNA sequences. *CABIOS* **3** (1987), 223-227.
1290. MENGERITSKY, G., AND SMITH, T. F. New analytical tool for analysis of splice site sequence determinants. *CABIOS* **5** (1989), 97-100.
1291. MENGERITSKY, G., AND TRIFONOV, G. Nucleotide sequence-directed mapping of the nucleosomes. *Nucl. Acids Res.* **11**(11) (1983), 3833-3851.
1292. MENTEN, L. E., KOHN, M. C., AND GARFINKEL, D. A convenient computer program for estimation of enzyme and metabolite concentrations in multienzyme systems. *Computers Biomed. Res.* **14** (1981), 91.
1293. MERKULOVA, T. I., SELEDTSOV, I. A., SOLOVYEV, V. V., PLISOV, S. Yu., AND NIKULINA, E. B. Identification of nucleotide sequences in human and animal genomes responsible for binding of glucocorticoid-receptor complexes and hormone regulation of transcription, (in) *Human Genome-90*, Bayev, A.E., (ed.), (Abstracts of the 1st National Conference, Pereslavl-Zalesky, Moscow), pp.123-124, 1990.
1294. MEWES, H. W., ELZANOWSKI, A., AND GEORGE, D. G. Protein sequence databases: database management, data structures and data access. *Biochem. Soc. Trans.* **17** (1989), 843-845.
1295. MEWES, H. W., AND TSUGITA, A. Additional protein sequence data banks, (in) *Computational Molecular Biology. Sources and Methods for Sequence Analysis*, Lesk, A.M., (ed.), (Oxford University Press, Oxford), pp.38-40, 1988.
1296. MEYERS, S. A simulator for regulatory genetics and its application to bacteriophage lambda. *Nucl. Acids Res.* **12**, 1 (1984), 1-9.
1297. MEYERS, S., AND FIEDLAND, P. Knowledge-based simulation of genetic regulation in bacteriophage lambda. *Nucl. Acids Res.* **12** (1984), 1-9.
1298. MICHEL, C. J. New statistical approach to discriminate between protein coding and non-coding regions in DNA sequences and its evaluation. *J. Theor. Biol.* **120** (1986), 223-236.
1299. MICHEL, C. J. A study of the purine/pyrimidine codon occurrence with a reduced centered variable and an evaluation compared to frequency statistic. *Math. Biosci.* **97** (1989), 161-178.
1300. MILLER, R. A. Preparing a medical informatics research grant proposal. *Computers Biomed. Res.* **22** (1989), 92-101.
1301. MILLER, W., BARR, J., AND RUDD, K. E. Improved algorithms for searching restriction maps. *CABIOS* **7**(4) (1991), 447-456.
1302. MILLER, W., AND MYERS, E. Sequence comparison with concave weighting functions. *Bull. Math. Biol.* **50**(2) (1988), 97-120.
1303. MILLER, W., AND MYERS, E. W. A file comparison program. *Software Practice and Experience* **15** (1985), 1025-1040.
1304. MILNE, R. Computer array interprets the human genome. *New Scientist* **122**(66) (1989), 36.
1305. MILOSAVLJEVIC, A., HASSLER, D., AND JURKA, J. Informed parsimonious inference of prototypical genetic sequences, (in) *Proc. Second Workshop on Computational Learning Theory*, preprint, 1989.
1306. MILOSAVLJEVIC, A. D. Categorization of macromolecular sequences by minimal length encoding, Ph.D. Thesis, University of California at Santa Cruz, 1991.
1307. MIL'SHINA, N. V. Discovering of potential "mRNA stability" signals in 3'-nontranslated regions mRNA of higher eukaryote nuclear genes, (in) *Human Genome-90*, Bayev, A.E., (ed.), (Abstracts of the 1st National Conference, Pereslavl-Zalesky, Moscow), pp.125-126, 1990.
1308. MIL'SHINA, N. V. Sequence signals which may be required for coordinate regulation of genes in mammals: project of database, (in) *Modelling and Computer Methods in Molecular Biology and Genetics*, Kolchanov, N.A., Ratner, V.A., (eds.), (Abstracts of the International Conference, Novosibirsk), p.215, 1990.

1309. MIL'SHINA, N. V. The computer prediction of "mRNA stability" signals on 3'-nontranslated regions of nuclear genes, (in) *Modelling and Computer Methods in Molecular Biology and Genetics*, Kolchanov, N.A., Ratner, V.A., (eds.), (Abstracts of the International Conference, Novosibirsk), pp.30-31, 1990.
1310. MINGHETTI, P. P., GIBBS, P. E. M., AND NORMAN, A. W. Computer analysis of 1, 25-dihydroxyvitamin D₃-receptor regulated promoters: Identification of a candidate D₃-response element. *Biochem. Biophys. Res. Commun.* **162** (1989), 869-875.
1311. MIRONOV, A., AND KISTER, A. RNA secondary structure formation during transcription. *J. Biomolec. Struct. Dyn.* **4** (1986), 1-9.
1312. MIRONOV, A. A., AND ALEXANDROV, N. N. Statistical method for rapid homology search. *Nucl. Acids Res.* **16** (1988), 5169-5173.
1313. MIRONOV, A. A., ALEXANDROV, N. N., LIUNOVSKAIA-GUROVA, L. V., AND KISTER, A. Computer programs for the analysis of nucleotide sequences (MALIK). *Mol. Biol.* (URSS) **21**:672-677, 1987.
1314. MITCHELL, E. M., ARTYMIUK, P. J., RICE, D. W., AND WILLETT, P. Use of techniques derived from graph theory to compare secondary structure motifs in protein. *J. Mol. Biol.* **212** (1989), 151-166.
1315. MIURA, R. M. , (ed.). *Some Mathematical Questions in Biology: DNA Sequence Analysis*, Lectures on Mathematics in the Life Sciences, (The American Mathematical Society, Providence, Rhode Island), 1986.
1316. MIYATA, T., AND HAYASHIDA, H. Extraordinary high evolutionary rate of pseudogenes: evidence for the presence of selective pressure against changes between synonymous codons. *Proc. Natl. Acad. Sci. USA* **78** (1981), 5739-5743.
1317. MIYATA, T., YASUNAGA, T., AND NISHIDA, T. Nucleotide sequence divergence and functional constraint in mRNA evolution. *Proc. Natl. Acad. Sci. USA* **77** (1980), 7328-7332.
1318. MIYAZAWA, S., AND JERNIGEN, R. L. Estimation of effective interresidue contact energies from protein crystal structures: quai-chemical approximation. *Macromolecules* **18** (1985), 534-552.
1319. MIZRAJI, E., AND NINIO, J. Graphical coding of nucleic acid sequences. *Biochimie* **67** (1985), 445-448.
1320. MOBERG, D., AND JOSEPHSON, J. Implementation note on diagnosing and fixing faults in theories, preprint, 1989.
1321. MODELEVSKY, J. L. Computer applications in recombinant DNA research. *Basic Life Science* **25** (1983), 17-34.
1322. MODELEVSKY, J. L. Computer applications in applied genetic engineering. *Adv. Appl. Microbiol.* **30** (1984), 169-195.
1323. MODELEVSKY, J. L., AND AKERS, T. G. 3-D multi-variate data display tool as a protein design aid. *CABIOS* **4** (1988), 308.
1324. MODELEVSKY, J. L., AND AKERS, T. G. Reiterative RNA folding and occupancy rate analysis for mRNA modelling. *CABIOS* **4** (1988), 161-166.
1325. MODIANO, G., BATTISTUZZI, G., AND MOTULSKY, A. Nonrandom patterns of codon usage and of nucleotide substitutions in human alpha- and beta-globin genes: An evolutionary strategy reducing the rate of mutations with drastic effects. *Proc. Natl. Acad. Sci. USA* **78** (1981), 1110-1114.
1326. MOODY, M., AND FRISTENSKY, B. Database bias and the identification of protein coding sequences. *DNA* **6** (1987), 493-496.
1327. MOON, J. B., AND HOWE, W. J. A fast algorithm for generating smooth molecular dot surface representations. *J. Mol. Graphics* **7** (1989), 109-112.
1328. MOORE, G. P., AND MOORE, A. R. The average spacing of restriction enzyme recognition sites in DNA. *J. Theor. Biol.* **98** (1982), 165-169.
1329. MOORE, G. W. Proof of the populous path algorithm for missing mutations in parsimony trees. *J. Theor. Biol.* **66** (1977), 95-106.
1330. MOOTE, N. J., CHENIER, P., MIKHAIL, M., AND DROUIN, J. A program for the graphic representation and manipulation of DNA sequences. *CABIOS* **3** (1987), 189-192.
1331. MORALES, L. B., GARDUNO-JUAREZ, R., AND ROMERO, D. Applications of simulated annealing to the multiple-minima problem in small peptides. *J. Biomol. Struct. Dyn.* **8(4)** (1991), 721-735.
1332. MOREAU, J., MARCAUD, L., MASCHAT, F., KEIZLAROVA-LEPESANT, J., LEPESANT, J., AND SCHERRER, K. A +t-rich linkers define functional domains in eukaryotic dna. *Nature* **295** (1982), 260-262.
1333. MORFFEW, A. J., AND TODD, S. J. P. The use of PROLOG as a protein querying language. *Computers and Chemistry* **10** (1986), 9-14.
1334. MORFFEW, A. J., TODD, S. J. P., AND SNELGROVE, M. J. The use of a relational data base for holding molecule data in a molecular graphics system. *Computers and Chemistry* **7** (1983), 9-16.
1335. MOROWITZ, H. J., AND SMITH, T. Report of the Matrix of Biological Knowledge Workshop, 1987.
1336. MORRIS, G. M. The matching of protein sequences using color intrasequence homology displays. *J. Mol. Graphics* **6** (1988), 135-142.
1337. MOSIMANN, J. E., SHAPIRO, M. B., MERRIL, C. R., BRADLEY, D. F., AND VINTON, J. E. Reconstruction of protein and nucleic acid sequences: IV. The algebra of free monoids and the fragmentation strategem. *Bull. Math. Biophys.* **28** (1966), 235-260.

1338. MOTT, R. Maximum-likelihood estimation of the statistical distribution of Smith-Waterman local sequence similarity scores. *Bull. Math. Biol.* **54**(1) (1992), 59-75.
1339. MOTT, R., KIRKWOOD, T. B. L., AND CURNOW, R. N. An accurate approximation of the length of the longest matching word between two random dna sequences. *Bull. Math. Biol.* **52** (1990), 773-784.
1340. MOTT, R. F., AND KIRKWOOD, T. B. L. STATSEARCH:a GCG-compatible program for assessing statistical significance during DNA and protein databank searches. *CABIOS* **6**(3) (1990), 293-295.
1341. MOTT, R. F., KIRKWOOD, T. B. L., AND CURNOW, R. N. A test for the statistical significance of DNA sequence similarities for application in databank searches. *CABIOS* **5** (2) (1989), 123-131.
1342. MOTT, R. F., KIRKWOOD, T. B. L., AND CURNOW, R. N. An accurate approximation to the distribution of the length of the longest matching word between two random DNA sequences. *Bull. Math. Biol.* **52** (1990), 773-784.
1343. MOTT, R. F., KIRKWOOD, T. B. L., AND CURNOW, R. N. Tests for the statistical significance of protein sequence similarities in data-bank searches. *Protein Engineering* **4**(2) (1990), 149-154.
1344. MOUCHIROUD, D., FICHANT, G., AND BERNARDI, G. Compositional compartmentalization and gene composition in the genome of vertebrates. *J. Mol. Evol.* **26** (1987), 198-204.
1345. MOUCHIROUD, D., AND GAUTIER, G. Codon usage changes and sequence dissimilarity between human and rat. *J. Mol. Evol.* **31** (1990), 81-91.
1346. MOUCHIROUD, D., GAUTIER, G., AND BERNARDI, G. The compositional distribution of coding sequences and DNA molecules in humans and murids. *J. Mol. Evol.* **27** (1988), 311-320.
1347. MOULT, J., AND JAMES, M. N. G. An algorithm for determining the conformation of polypeptide segments in proteins by systematic search. *Proteins* **1** (1986), 146-163.
1348. MOUNT, D. W. Modeling RNA structure. *Bio/Technology* **2** (1984), 791-795.
1349. MOUNT, D. W. Computer analysis of sequence, structure and function of biological macromolecules. *Biotechniques* **3** (1985), 102-112.
1350. MOUNT, S. M. A catalogue of splice junction sequences. *Nucl. Acids Res.* **10** (1982), 459-472.
1351. MRAZEK, J., AND KYPR, J. Computer program Jamseq combining statistical and stereochemical rules for the prediction of protein secondary structure. *CABIOS* **4** (1988), 297-302.
1352. MRAZEK, J., AND KYPR, J. GLOBIC: a very fast microcomputer program for fingerprinting, characterization and comparison of long nucleotide sequences. *CABIOS* **8**(1) (1992), 29-34.
1353. MUKHERJEE, A. Hardware algorithms for determining similarity between two strings. *IEEE Transactions on Computers* **38** (1989), 600-603.
1354. MULLER, V., AND FITCH, W. M. Evolutionary selection for perfect hairpin structures in viral DNAs. *Nature* **298** (1982), 582-585.
1355. MULLIGAN, M. E., HAWLEY, D. K., ENTRIKEN, R., AND MCCLURE, W. R. *Escherichia coli* promoter sequences predict in vitro RNA polymerase selectivity. *Nucl. Acids Res.* **12** (1984), 789-800.
1356. MULLIGAN, M. E., AND MCCLURE, W. R. Analysis of the occurrence of promoter-sites in DNA. *Nucl. Acids Res.* **14** (1986), 109-126.
1357. MURATA, M. Three-way Needleman-Wunsch algorithm, (in) *Methods in Enzymology-Molecular Evolution: Computer Analysis of Protein and Nucleic Acid Sequences*, Vol.183, Doolittle, R.F., (ed.), (Academic Press, NY), pp.365-375, 1990.
1358. MURATA, M., RICHARDSON, J. S., AND SUSSMAN, J. L. Simultaneous comparison of three protein sequences. *Proc. Natl. Acad. Sci.* **82** (1985), 3073-3077.
1359. MURRAY, E. E., COTZER, J., AND EBERLE, M. Codon usage in plant genes. *Nucl. Acids Res.* **17** (1989), 477-498.
1360. MUZARD, G., THEVENY, B., AND REVET, B. Electron microscopy mapping of pBR322 DNA curvature. Comparison with theoretical models. *EMBO* **9**(4) (1990), 1289-1298.
1361. MYERS, E. W. An O(ND) difference algorithm and its variations. *Algorithmica* **1**(2) (1986), 251-266.
1362. MYERS, E. W. Incremental alignment algorithms and their applications. *SIAM J. Computing* (1989). to be published.
1363. MYERS, E. W., AND HUANG, X. An $O(N/u/s - 22/s_0/d\log/N)$ restriction map comparison and search algorithm, *Bull. Math. Biol.*, in press.
1364. MYERS, E. W., AND MILLER, W. Optimal alignments in linear space. *CABIOS* **4**(1) (1988), 11-17.
1365. MYERS, E. W., AND MILLER, W. Approximate matching of regular expressions. *Bull. Math. Biol.* **51**(1) (1989), 5-37.
1366. MYERS, E. W., AND MOUNT, D. Computer program for the IBM personal computer that searches for approximate matches of short oligonucleotide sequences in long target DNA sequences. *Nucl. Acids Res.* **14**(1) (1986), 501-508.
1367. NAGANO, K. Triplet information in helix prediction applied to the analysis of super-secondary structure. *J. Mol. Biol.* **109** (1977), 251-257.
1368. NAGANO, K. A computer method of finding supersecondary structure, (in) *Methods in Protein Sequence Analysis*, Jornvall, H., Hoog, J.-O., and Gustavsson, A.-M., (ed.), (Birkhauser Verlag, Boston, MA), pp.333-341, 1991.
1369. NAKAI, K., AND KIDERA, A. KANEHISA, M. Cluster analysis of amino acid indices for prediction of protein structure and function. *Protein Engineering* **2**(2) (1988), 93-100.

1370. NAKASHIMA, H., NISHIKAWA, K., AND OOI, T. The folding type of a protein is relevant to the amino acid composition. *J. Biochem.* **99** (1986), 153-162.
1371. NAKATA, K., KANEHISA, M. I., AND DELISI, C. Prediction of splice junctions in mRNA sequences. *Nucl. Acids Res.* **13** (1985), 5327-5340.
1372. NAKATA, K., KANEHISA, M. I., AND MAIZEL, J. V. Jr. Discriminant analysis of promoters regions in *Escherichia coli* sequences. *CABIOS* **4** (1988), 367-371.
1373. NAKATA, K., AND MAIZEL, J. V. Prediction of operator-binding protein by discriminant analysis. *Gene Anal. Techn.* **6** (1989), 111-119.
1374. NAORA, H., AND DEACON, N. Relationships between the total size of exons and introns in protein-coding genes of higher eukaryotes. *Proc. Natl. Acad. Sci. USA* **79** (1982), 6196-6200.
1375. NAORA, H., MIYAHARA, K., AND CURNOW, R. Origin of noncoding DNA sequences: Molecular fossils of genome evolution. *Proc. Natl. Acad. Sci. USA* **84** (1987), 6195-6199.
1376. NAQVI, S., AND TSUR, S. *A Logical Language for Data and Knowledge Bases*, (Computer Science Press, NY), 1989.
1377. NARAYANA, S. V. L., AND ARGOS, P. Residue contacts in proteins structures and implications for protein folding. *Int. J. Pept. Protein Res.* **24** (1984), 25-39.
1378. NASSIMI, D., AND SAHNI, S. Parallel permutation and sorting algorithms and a new generalized connection network. *J. ACM* **29** (1982), 642-667.
1379. NAYLOR, S., FINDEIS, A. F., GIBSON, B. W., AND WILLIAMS, D. H. An approach toward the complete FAB analysis of enzymic digests of peptides and proteins. *J. Am. Chem. Soc.* **108** (1986), 6359-6363.
1380. NEDELMAN, J., HEAGERTY, P., AND LAWRENCE, C. Quantitative PCR with internal controls. *CABIOS* **8(1)** (1992), 65-70.
1381. NEEDLEMAN, S. B., AND WUNSCH, C. D. A general method applicable to the search for similarities in the amino acid sequence of 2 proteins. *J. Mol. Biol.* **48** (1970), 443-453.
1382. NEI, M. *Molecular Evolutionary Genetics*. (Columbia University Press, NY), 1987.
1383. NEI, M., AND GOJOBORI, T. Simple methods for estimating the number of synonymous and nonsynonymous nucleotide substitutions. *Mol. Biol. Evol.* **3** (1986), 418-426.
1384. NEI, M., AND MILLER, J. C. A simple method for estimating average number of nucleotide substitutions within and between populations from restriction data. *Genetics* **125** (1990), 873-879.
1385. NEI, M., STEPHENS, J. C., AND SAITOU, N. Methods for computing the standard errors of branching points in an evolutionary tree and their application to molecular data from humans and apes. *Mol. Biol. Evol.* **2(1)** (1985), 66-85.
1386. NEI, M., AND TATENO, Y. Nonrandom amino acid substitution and estimation of the number of nucleotide substitutions in evolution. *J. Mol. Evol.* **11** (1978), 333-347.
1387. NEI, M., AND TATENO, Y. Augmentation algorithm: a reply to holmquist. *J. Mol. Evol.* **13** (1979), 167-171.
1388. NEIDLE, S., AND JENKINS, T. C. Molecular modeling to study DNA intercalation by anti-tumor drugs, (in) *Methods in Enzymology-Molecular Design and Modeling: Concepts and Applications*, Part B, Vol.203, Langone, J.J., (ed.), (Academic Press, NY), pp.433-458, 1991.
1389. New Announcements from the Human Genome Program at NIH, (in) *NIH Guide for Grants and Contracts*, Vol.18, No.25, July 21-28, 1989.
1390. NICEWANDER, W. A., AND PRICE, J. M. Reliability of measurement and the power of statistical test: Some new results. *Psych. Bull.* **94** (1983), 524-533.
1391. NICHOLAS, H. B. Jr., AND MCCLAIN, W. H. An algorithm for discriminating sequences and its applications to yeast transfer RNA. *CABIOS* **3(3)** (1987), 177-181.
1392. NIERMANN, T., AND KIRSCHNER, K. Use of homologous sequences to improve protein secondary structure prediction, (in) *Methods in Enzymology-Molecular design and Modeling: Concepts and Applications*, Part A, Vol.202, Langone, J.J., (ed.), (Academic Press, NY), pp.45-59, 1991.
1393. NILGES, M., CLORE, G. M., AND GRONENBORN, A. M. H-NMR stereospecific assignments by conformational data-base searches. *Biopolymers* **29** (1990), 813-822.
1394. NINIO, J. Properties of nucleic acid representations I. Topology. *Biochimie* **53** (1971), 485-494.
1395. NINIO, J., AND MIZRAJI, E. String analysis and energy minimization in the partition of DNA sequences. *J. Mol. Biol.* **207** (1989), 585-596.
1396. NISHIKAWA, K. Assessment of secondary-structure prediction of proteins. Comparison of computerized Chou-Fasman method with others. *Biochim. Biophys. Acta* **748** (1983), 285-299.
1397. NISHIKAWA, K. Prediction of the secondary structure of proteins: a method of using homology to proteins of known structure. *Protein, Nucleic Acid and Enzyme* **29** (1986), 79-89.
1398. NISHIKAWA, K., KUBOTA, Y., AND OOI, T. Classification of proteins into groups based on amino acid composition and other characters. I. Angular distribution. *J. Biochem.* **94** (1983), 981-995.
1399. NISHIKAWA, K., KUBOTA, Y., AND OOI, T. Classification of proteins into groups based on amino acid composition and other characters. II. Grouping into four types. *J. Biochem.* **94** (1983), 997-1007.
1400. NISHIKAWA, K., NAKASHIMA, H., KANEHISA, M., AND OOI, T. Detection of weak sequence homology of proteins for tertiary structure prediction. *Protein Seq. Data Anal.* **1** (1987), 107-116.

1401. NISHIKAWA, K., AND NOGUCHI, T. Predicting protein secondary structure based on amino acid sequence, (in) *Methods in Enzymology-Molecular design and Modeling: Concepts and Applications*, Part A, Vol.202, Langone, J.J., (ed.), (Academic Press, NY), pp.31-44, 1991.
1402. NISHIKAWA, K., AND OOI, T. Correlation of amino acid composition of a protein with its structural and biological characters. *J. Biochem.* 91 (1982), 1821-1824.
1403. NISHIKAWA, K., AND OOI, T. Radial locations of amino acid residues in a globular protein: correlation with the sequences. *J. Biochem.* 100 (1986), 1043-1047.
1404. NOETZEL, A., AND SELKOW, S. M. An analysis of the general tree-editing problem, (in) *Time Warps, String Edits, and Macromolecules: The Theory and Practice of Sequence Comparison*, Sankoff, D., Kruskal, J.B., (eds.), (Addison-Wesley, Reading, MA), 1983.
1405. NOLAN, G. P., MAINA, C. V., AND SZALAY, A. A. Plasmid mapping computer program. *Nucl. Acids Res.* 12 (1984), 717-729.
1406. NORMAN, C. Bush budget highlights R.D. *Science* 247 (1990), 517-519.
1407. NORTH, P. A., AND MALCOLM, A. D. B. Prediction of restriction sites from protein sequences: a computer program. *Biochem. Soc. Trans.* 13 (1985), 789-791.
1408. NOVOTNY, J. Matrix program to analyze primary structure homology. *Nucl. Acids Res.* 10 (1982), 127-131.
1409. NOVOTNY, J., AND AUFRAY, C. A program for prediction of protein secondary structure from nucleotide sequence data: application to histocompatibility antigens. *Nucl. Acids Res.* 12 (1984), 243-255.
1410. NOVOTNY, J., BRUCCOLERI, R. E., AND KARPLUS, M. An analysis of incorrectly folded models-implications for structure predictions. *J. Mol. Biol.* 177 (1984), 787-818.
1411. NOZAKI, Y., AND TANFORD, C. The solubility of amino acids and two glycine peptides in aqueous ethanol and dioxane solutions: establishment of a hydrophobicity scale. *J. Biol. Chem.* 246 (1971), 2211-2217.
1412. NUSSINOV, R. Some rules for ordering nucleotides in DNA. *Nucl. Acids Res.* 8 (1980), 4545-4562.
1413. NUSSINOV, R. Strong adenine clustering in nucleotide sequences. *J. Theor. Biol.* 85 (1980), 285-291.
1414. NUSSINOV, R. Eukaryotic dinucleotide preference rules and their implications for degenerate codon choice. *J. Mol. Biol.* 149 (1981), 125-131.
1415. NUSSINOV, R. Nearest neighbor nucleotide patterns. Structural and biological implications. *J. Biol. Chem.* 256 (1981), 8458-8462.
1416. NUSSINOV, R. The universal dinucleotide asymmetry rules in DNA and the amino acid codon choice. *J. Mol. Evol.* 17 (1981), 237-244.
1417. NUSSINOV, R. Some indications for inverse DNA duplication. *J. Theor. Biol.* 95 (1982), 783-791.
1418. NUSSINOV, R. Efficient algorithms for searching for exact repetition of nucleotide sequences. *J. Mol. Evol.* 19 (1983), 283-285.
1419. NUSSINOV, R. An efficient code searching for sequence homology and DNA duplication. *J. Theor. Biol.* 100 (1983), 319-328.
1420. NUSSINOV, R. Doublet frequencies in evolutionarily distinct groups. *Nucl. Acids Res.* 12 (1984), 1749-1763.
1421. NUSSINOV, R. Promoter helical structure variation at the *Escherichia coli* polymerase interaction sites. *J. Biol. Chem.* 259 (1984), 6798-6805.
1422. NUSSINOV, R. Strong doublet preferences in nucleotide sequence and DNA geometry. *J. Mol. Evol.* 20 (1984), 111-119.
1423. NUSSINOV, R. Structural features are as important as sequence homologies in drosophila heat shock upstream regions. *J. Mol. Evol.* 20 (1984), 106-110.
1424. NUSSINOV, R. Large helical conformational deviations from ideal B-DNA and prokaryotic regulatory sites. *J. Theor. Biol.* 115 (1985), 179-189.
1425. NUSSINOV, R. The predicted presence of large helical structure variation in yeast his-4 upstream region is correlated with general amino acid control on the cyc-1 gene. *J. Biomol. Struct. Dyn.* 3 (1985), 349-361.
1426. NUSSINOV, R. Structural wrinkles and the genomic regulatory sites of eukaryotes. *J. Mol. Evol.* 22 (1985), 150-159.
1427. NUSSINOV, R. Compilation of eukaryotic sequences around transcription initiation sites. *J. Theor. Biol.* 120 (1986), 479-487.
1428. NUSSINOV, R. Sequence signals which may be required for efficient formation of mRNA 3' termini. *Nucl. Acids Res.* 14 (1986), 3557-3571.
1429. NUSSINOV, R. Some guidelines for identification of recognition sequences: regulatory sequences frequently contain T(GTG)/CAC(A), TGA/TCA and (T)CTC/GAG(A). *Biochim. Biophys. Acta* 866 (1986), 93-108.
1430. NUSSINOV, R. TGTG, G clustering and other signals near non-mammalian vertebrate mRNA 3' termini: some implications. *J. Biomol. Struct. Dyn.* 3 (1986), 1145-1153.
1431. NUSSINOV, R. (A)GGG(A), (A)CCC(A) and other potential 3' splice signals in primate nuclear pre-mRNA sequences. *Biochim. Biophys. Acta* 910 (1987), 261-270.
1432. NUSSINOV, R. Asymmetry in the distributions of the four nucleotides at mRNA initiation and 3' termini sites: some geometrical implications. *Biochim. Biophys. Acta* 908 (1987), 143-149.
1433. NUSSINOV, R. Conserved putative signals in 3' intron junctions in rodents. *J. Biomol. Struct. Dyn.* 4 (1987), 1051-1064.

1434. NUSSINOV, R. Conserved quartets near 5' intron junctions in primate nuclear pre-mRNA. *J. Theor. Biol.* **133** (1987), 73-84.
1435. NUSSINOV, R. Nucleotide quartets in the vicinity of eukaryotic transcriptional initiation sites: some DNA and chromatin structure implications. *DNA* **6** (1987), 13-22.
1436. NUSSINOV, R. Theoretical molecular biology: prospectives and perspectives. *J. Theor. Biol.* **125** (1987), 219-235.
1437. NUSSINOV, R. Putative elements in the vicinity of viral transcription initiation sites. *Int. J. Biochem.* **20** (1988), 721-730.
1438. NUSSINOV, R. Conserved signals around the 5' splice sites in eukaryotic nuclear precursor mRNAs: G-runs are frequent in the introns and C in the exons near both 5' and 3' splice sites. *J. Biomol. Struct. Dyn.* **6** (1989), 985-1000.
1439. NUSSINOV, R. The ordering of nucleotides in DNA: computational problems in molecular biology. *Comput. Biol. Med.* **19** (1989), 269-281.
1440. NUSSINOV, R. General nearest neighbor preferences in G/C oligomers interrupted by A/T: Correlation with DNA structure. *J. Biomol. Struct. Dyn.* **8** (1990), 399-411.
1441. NUSSINOV, R., AND JACOBSON, A. B. Fast algorithm for predicting the secondary structure of single-stranded RNA. *Proc. Natl. Acad. Sci. USA* **77** (1980), 6309-6313.
1442. NUSSINOV, R., AND LENNON, G. G. Periodic structurally similar oligomers are found on one side of the axes of symmetry in the lac, trp and gal operators. *J. Biomol. Struct. Dyn.* **2** (1984), 387-395.
1443. NUSSINOV, R., OWENS, J., AND MAIZEL, J. V. Jr. Sequence signals in eukaryotic upstream regions. *Biochim. Biophys. Acta* **866** (1986), 109-119.
1444. NUSSINOV, R., PIECZENIK, G., GRIGGS, J., AND KLEITMAN, D. Algorithms for loop matchings. *SIAM J. Appl. Math.* **35** (1978), 68-82.
1445. NUSSINOV, R., SARAI, A., SMYTHERS, G., AND JERNIGAN, R. Sequence context of oligomer tracts in eukaryotic DNA: biological and conformational implications. *J. Biomol. Struct. Dyn.* **6** (1988), 543-562.
1446. NUSSINOV, R., SARAI, A., SMYTHERS, G., AND JERNIGAN, R. Distinct patterns in homooligomer tracts sequence context in prokaryotic and eukaryotic DNA. *Biochim. Biophys. Acta* **100** (1989), 329-338.
1447. NUSSINOV, R., SARAI, A., SMYTHERS, G., AND WANG, D. Strong patterns in homooligomer tracts occurrences in non-coding and in potential regulatory sites in eukaryotic genomes. *J. Biomol. Struct. Dyn.* **7** (1989), 707-722.
1448. NUSSINOV, R., SARAI, A., WANG, D., AND JERNIGAN, R. L. Sequence context of homooligomer tracts in eukaryotic genomes: some DNA conformational implications, Olson W.K., Sarma M.K., Sarma R.H., Sundaralingam M., (Adenine Press, NY), 1988.
1449. NUSSINOV, R., SHAPIRO, B., LIPKIN, L., AND MAIZEL, J. DNAase I hypersensitive sites may be correlated with genomic regions of large structural variations. *J. Mol. Biol.* **177** (1984), 591-607.
1450. NUSSINOV, R., SHAPIRO, B., LIPKIN, L., AND MAIZEL, J. Enhancer elements share local homologous twist-angle variations with a helical periodicity. *Biochim. Biophys. Acta* **783** (1984), 246-257.
1451. NUSSINOV, R., AND TINOCO, I. Sequential folding of a messenger RNA molecule. *J. Mol. Biol.* **151** (1981), 519-533.
1452. NUSSINOV, R., AND TINOCO, I. Small changes in free energy assignments for unpaired bases do not affect predicted secondary structures in single stranded RNA. *Nucl. Acids Res.* **10** (1982), 341-350.
1453. NUSSINOV, R., TINOCO, I. Jr., AND JACOBSON, A. B. Secondary structure model for the complete simian virus 40 late precursor mRNA. *Nucl. Acids Res.* **10** (1982), 351-363.
1454. NUSSINOV, R., AND WOLFSON, H. J. Efficient detection of three-dimensional structural motifs in biological macromolecules by computer vision techniques. *Proc. Natl. Acad. Sci. USA* **88** (1991), 10495-10499.
1455. OCHS, R. S., AND CONROW, K. A computerized metabolic map. *J. Chem. Inf. Comput. Sci.* **31** (1991), 132-137.
1456. O'DONOGHUE, S. I. Structural interpretation of fluorescence resonance-energy transfer measurements. *CABIOS* **7(4)** (1991), 471-477.
1457. OERTER, K. E., MUNSON, P. J., McBRIDE, W. O., AND RODBARD, D. Computerized estimation of size of nucleic acid fragments using the four-parameter logistic model. *Analytical Biochemistry* **189(2)** (1990), 235-243.
1458. O'FERRELL, D., JENSON, E., AND LINSLEY, J. A study of the movement of DNA molecules in field inversion gel electrophoresis, (in) *Proc. 1987 AIChE Annual Meeting*, preprint, 1987.
1459. OHNO, S. Codon preference is but an illusion created by the construction principle of coding sequences. *Proc. Natl. Acad. Sci. USA* **85** (1988), 4378-4382.
1460. OHNO, S. Universal rule for coding sequence construction:TA/CG deficiency-TG/CT excess. *Proc. Natl. Acad. Sci. USA* **85** (1988), 9630-9634.
1461. OHSHIMA, Y., AND GOTOH, Y. Signals for the selection of a splice site in pre-mRNA. Computer analysis of splice junctions sequences and like sequences. *J. Mol. Biol.* **197** (1987), 247-259.
1462. OHTA, T. Population genetics theory of concerted evolution and its application to the immunoglobulin V gene tree. *J. Mol. Evol.* **20** (1985), 274-280.
1463. OLAFSON, B. D., MARUSIN, J. R., AND ARY, M. L. Macromolecule structure determination using NMR data and molecular simulation techniques. *American Biotechnology Laboratory* **8(8)** (1990), 18.

1464. OLIVER, J. L., MARIN, A., AND MARTINEZ-ZAPATER, J. M. Chloroplast genes transferred to the nuclear plant genome have adjusted to nuclear base composition and codon usage. *Nucl. Acids Res.* **18** (1990), 65-73.
1465. OLSON, M., HOOD, L., CANTOR, C., AND BOTSTEIN, D. A common language for physical mapping of the human genome. *Science* **234** (1985), 1434-1435.
1466. OLSON, S., KRAUS, W., AND CORTELYOU, M. Sequence analysis with full features table support. *Macromolecules, Genes, and Computers Symposium*, Waterville Valley, NH, August 13-18, 1989. (abstract only).
1467. OLSON, W. K., AND SRINIVASAN, A. R. The translation of DNA primary base sequence in three-dimensional structure. *CABIOS* **4**(1) (1988), 133-142.
1468. OLSON, W. K., AND ZHANG, P. Computer simulation of DNA supercoiling, (in) *Methods in Enzymology-Molecular Design and Modeling: Concepts and Applications*, Part B, Vol.203, Langone, J.J., (ed.), (Academic Press, NY), pp.403-432, 1991.
1469. O'NEILL, M. C. Consensus methods for finding and ranking DNA binding sites. Application to *Escherichia coli* promoters. *J. Mol. Biol.* **207** (1989), 301-310.
1470. O'NEILL, M. C. *Escherichia coli* promoters. I. Consensus as it relates to spacing class, specificity, repeat substructure, and three-dimensional organization. *J. Biol. Chem.* **264** (1989), 5522-5530.
1471. O'NEILL, M. C., AND CHIAFARI, F. *Escherichia coli* promoters. II. A spacing-class dependent promoter search protocol. *J. Biol. Chem.* **264** (1989), 5531-5534.
1472. OOI, T., AND TAKAHASHI, S. Correlation between primary structure and tertiary structure of proteins. *Protein, Nucleic Acid and Enzyme* **29** (1986), 69-78.
1473. OOI, T., AND TAKANAMI, M. A computer method for construction of secondary structure from polynucleotide sequence. Possible structure of the bacterial replication origin. *Biochim. Biophys. Acta* **65** (1981), 221-229.
1474. ORCUTT, B. C., AND BARKER, W. C. Searching the protein database. *Bull. Math. Biol.* **46** (1984), 545-552.
1475. ORCUTT, B. C., GEORGE, D. G., AND DAYHOFF, M. O. Protein and nucleic acid sequence database systems. *Ann. Rev. Biophys. Bioeng.* **12** (1983), 419-441.
1476. ORCUTT, B. C., GEORGE, D. G., FREDRICKSON, J. A., AND DAYHOFF, M. O. Nucleic acid sequence database computer system. *Nucl. Acids Res.* **10**(1) (1982), 157-174.
1477. ORENGO, C. A., AND TAYLOR, W. R. A rapid method of protein structure alignment. *J. Theor. Biol.* **147** (1990), 517-551.
1478. ORLOWSKI, M., AND PACHTER, M. An algorithm for the determination of a longest increasing subsequence in a sequence. *Computers and Mathematics with Applications* **17** (1989), 1073-1075.
1479. OSAWA, S., OHAMA, T., YAMAO, F., MUTO, A., JUKES, T., OZEKI, H., AND UMESONO, K. Directional mutation pressure and transfer RNA in choice of the third nucleotide of synonymous two-codon sets. *Proc. Natl. Acad. Sci. USA* **85** (1988), 1124-1128.
1480. OSTELL, J. GenInfo ASN.1 Syntax: Sequences. Tech. Rep. 1, National Center for Biotechnology Information, National Library of Medicine, 1990.
1481. OSTELL, J. GenInfo Backbone Database Overview. Tech. Rep. 2, National Center for Biotechnology Information, National Library of Medicine, 1990.
1482. OSTERBURG, G., GLATTING, K. H., BUCHERT, J., AND WOLTERS, J. A fast method for arranging DNA sequence fragments. *Computer Programs in Biomedicine* **16** (1983), 61-70.
1483. OSTERBURG, G., GLATTING, K. H., AND SOMMER, R. Computer programs for the analysis and the management of DNA sequences. *Nucl. Acids Res.* **10** (1982), 207-216.
1484. OSTERBURG, G., AND SOMMER, R. Computer support of DNA sequence analysis. *Comput. Programs Biomed.* **35** (1981), 68-82.
1485. OTAKA, E., AND OOI, T. Examination of protein sequence homologies: V. New perspectives on evolution between bacterial and chloroplast-typeferredoxins inferred from sequence evidence. *J. Mol. Evol.* **29** (1989), 246-254.
1486. OTAKA, E., OOI, T., KUMAZAKI, T., AND ITOH, T. Examination of protein sequence homologies: I. Eleven *Escherichia coli* L7/L12-type ribosomal "A" protein sequences from eubacteria and chloroplast. *J. Mol. Evol.* **21** (1985), 339-345.
1487. OTSUKA, J., AND KUNISAVA, T. Characteristic base sequence patterns of promoter and terminator sites in phiX174 and fd phage DNAs. *J. Theor. Biol.* **97** (1982), 415-436.
1488. OVERBEEK, R., PRICE, M., AND TSUR, S. Automated interpretation of genetic sequencing gels, preprint, 1990.
1489. OWENS, J., CHATTERJEE, D., NUSSINOV, R., KONOPKA, A. K., AND MAIZEL, J. V. Jr. A fixed-point alignment technique for detection of recurrent and common sequences motifs associated with biological features. *CABIOS* **4**(1) (1988), 73-77.
1490. PABO, C. O. New generation databases for molecular biology. *Nature* **327** (1987), 467.
1491. PABO, C. O., AND SUCHANEK, E. G. Computer-aided model-building strategies for protein design. *Biochemistry* **25** (1986), 5987-5991.
1492. PAETKAU, V. A new rule for analyzing homologous coding sequences in DNA. *Nucl. Acids Res.* **12**(1) (1984), 159-173.

1493. PALAU, J., AND ARGOS, P. Protein secondary structure. *Int. J. Pept. Protein Res.* **19** (1982), 394-401.
1494. PALMER, T. C., HAUSHEER, F. H., AND SAXE, J. D. Applications of ray tracing in molecular graphics. *J. Mol. Graphics* **7** (1989), 160.
1495. PAOLELLA, G., AND RUSSO, T. A microcomputer program for the identification of tRNA genes. *Comput. Appl. Biosci.* **1** (1985), 149-151.
1496. PAPANICOLAOU, C., GOUY, M., AND NINIO, J. An energy model that predicts the correct folding of both the tRNA and the 5S RNA molecules. *Nucl. Acids Res.* **12**(1) (1984), 31-44.
1497. PARKER, J. M. R., GUO, D., AND HODGES, R. S. New hydrophilicity scale derived from high-performance liquid chromatography peptide retention data: correlation of predicted surface residues with antigenicity and X-ray-derived accessible sites. *Biochemistry* **25** (1986), 5425-5431.
1498. PASCARELLA, S., AND BOSSA, F. CLEAVAGE: a microcomputer program for predicting signal sequence cleavage sites. *CABIOS* **5**(1) (1989), 53-54.
1499. PASTOR, J. A., KOILE, K., AND OVERTON, G. Using analogy to predict functional regions on genes, (in) *Proceedings of the 24th Hawaii International Conference on System Science*, (Koloa, HA), Vol.I, pp.615-625, 1991.
1500. PASTORE, A., ATKINSON, R. A., SAUDEK, V., AND WILLIAMS, R. J. P. Topological mirror images in protein structure computation: An underestimated problem. *Proteins: Structure, Function, and Genetics* **10** (1991), 22-32.
1501. PATARCA, R., DORTA, B., AND RAMIREZ, J. L. Creation of a database for sequences of ribosomal nucleic acids and detection of conserved restriction endonucleases sites through computerized processing. *Nucl. Acids Res.* **10** (1982), 175-182.
1502. PATTHY, L. Intron-dependent evolution: preferred types of exons and introns. *FEBS Lett.* **214** (1984), 1-7.
1503. PATTHY, L. Detecting homology of distantly related proteins with consensus sequences. *J. Mol. Biol.* **198** (1987), 567-577.
1504. PATTON, P. C. Performance Limits for Parallel Processors, (in) *Parallel Supercomputing: Methods, Algorithms and Applications*, Carey, G.F., (ed.), (Wiley, NY), 1989.
1505. PAULETTI, D., SIMMONDS, R., DREESMAN, G. R., AND KENNEDY, R. C. Application of a modified computer algorithm in determining potential antigenic determinants associated with the AIDS virus glycoprotein. *Anal. Biochem.* **151** (1985), 540-546.
1506. PEACOCK, D., AND BOULTER, D. Use of amino acid sequence in phylogeny and evaluation of methods using computer simulations. *J. Mol. Biol.* **95** (1975), 513-527.
1507. PEARSON, P. K. Fast hashing of variable length text strings. *CACM* **33** (1990), 677-680.
1508. PEARSON, W. R. Automatic construction of restriction site maps. *Nucl. Acids Res.* **10** (1982), 217-227.
1509. PEARSON, W. R. Rapid and sensitive sequence comparison with FASTP and FASTA, (in) *Methods in Enzymology, Molecular Evolution: Computer Analysis of Protein and Nucleic Acid Sequences*, Vol.**183**, Doolittle, R.F., (ed.), (Academic Press, NY), pp.63-98, 1990.
1510. PEARSON, W. R., AND LIPMAN, D. J. Improved tools for biological sequence comparison. *Proc. Natl. Acad. Sci.* **85** (1988), 2444-2448.
1511. PELTOLA, H., SODERLUND, H., TARHIO, J., AND UKKONEN, E. Algorithms for some string matching problems arising in molecular genetics. *Information Processing* **83** (1983), 53-64.
1512. PELTOLA, H., SODERLUND, H., AND UKKONEN, E. SEQQAID: A DNA sequence assembling program based upon a mathematical model. *Nucl. Acids Res.* **12**(1) (1984), 307-321.
1513. PELTOLA, H., SODERLUND, H., AND UKKONEN, E. Algorithms for the search of amino acid patterns in nucleic acid sequences. *Nucl. Acids Res.* **14** (1986), 99-107.
1514. PENNY, D., AND HENDY, M. D. Estimating the reliability of evolutionary trees. *Mol. Biol. Evol.* **3** (1986), 403-417.
1515. PENNY, D., AND HENDY, M. D. Turbo tree: a fast algorithm for minimal trees. *CABIOS* **3**(3) (1987), 183-187.
1516. PENOTTI, F. E. Human DNA TATA boxes and transcription initiation sites. A statistical study. *J. Mol. Biol.* **213** (1990), 37-52.
1517. PEPKE, E., HASENFELD, A., LEVENE, S., CANTOR, C. R., AND LIM, H. A. Three dimensional simulation and graphical visualization of pulsed field gel electrophoresis, FSU-SCRI-91-104, August 1991.
1518. PERICOLAS, W. L., WANG, Y., AND THOMAS, G. A. Some rules for predicting the base-sequence dependence of DNA conformation. *Proc. Natl. Acad. Sci. USA* **85** (1988), 2579-2583.
1519. PERLWITZ, M. D., BURKS, C., AND WATERMAN, M. S. Pattern analysis of the genetic code. *Adv. Applied Math.* **9** (1988), 7-21.
1520. PERRELLA, F. W. EZ-FIT: A practical curve-fitting microcomputer program for the analysis of enzyme kinetic data on IBM PC compatible computers. *Analytical Biochemistry* **174** (1988), 437-447.
1521. PESOLE, G., ATTIMONELLI, M., AND LIUNI, S. A backtranslation method based on codon usage strategy. *Nucl. Acids Res.* **16** (1988), 1715-1728.
1522. PETERSEN, G. B., STOCKWELL, P. A., AND HILL, D. F. Messenger RNA recognition in *Escherichia coli*: a possible second site of interaction with 16S ribosomal RNA. *EMBO J.* **7** (1988), 3957-3962.

1523. PETERSON, S. B., BOHR, H., BOHR, J., BRUNAK, S., COTTERILL, R. M. J., FREDHOLM, H., AND LAUTRUP, B. Training neural networks to analyse biological sequences. *Trends in Biotechnology* 8(11) (1990), 304-308.
1524. PETRELLI, P. An algorithm for reconstructing protein sequences. *Int. J. Pept. Protein Res.* 25 (1985), 85-88.
1525. PEVZNER, P. A. 1-tuple DNA sequencing: computer analysis. *J. Biomol. Struct. Dyn.* 7 (1989), 63-73.
1526. PEVZNER, P. A., BORODOVSKY, M. Yu., AND MIRONOV, A. A. Linguistics of nucleotide sequences I: The significance of deviations from mean statistical characteristics and prediction of the frequencies of occurrence of words. *J. Biomol. Struct. Dyn.* 6 (1989), 1013-1026.
1527. PEVZNER, P. A., BORODOVSKY, M. Yu., AND MIRONOV, A. A. Linguistics of nucleotide sequences II: Stationary words in genetic texts and zonal structure of DNA. *J. Biomol. Struct. Dyn.* 6 (1989), 1027-1038.
1528. PEVZNER, P. A., AND KOZHUKHIN, C. G. Genome inhomogeneity is determined mainly by WW and SS dinucleotides, (in) *Modelling and Computer Methods in Molecular Biology and Genetics*, Kolchanov, N.A., Ratner, V.A., (eds.), (Abstracts of the International Conference, Novosibirsk), pp.92-93, 1990.
1529. PFEIFFER, F., AND GILBERT, W. A. VecBase: a cloning vector sequence data base. *Protein Seq. Data Anal.* 1 (1988), 269-280.
1530. PHILIPSON, L. The DNA data libraries. *Nature* 332 (1988), 676.
1531. PHILLIPS, G. J., ARNOLD, J., AND IVARIE, R. Mono-through hexanucleotide composition of the *Escherichia coli* genome: a Markov chain analysis. *Nucl. Acids Res.* 15 (1987), 2611-2626.
1532. PHILLIPS, G. J., ARNOLD, J., AND IVARIE, R. The effect of codon usage on the oligonucleotide composition of the *E. coli* genome and identification of over-and under represented sequences by Markov chain analysis. *Nucl. Acids Res.* 15 (1987), 2627-2638.
1533. PICARD, B., DENIS, J. B., DECOUX, G., AND GOULET, P. Computer-assisted statistical analyses of enzyme and ribosomal DNA electrophoretic polymorphism in *Yersinia*. *Electrophoresis* 12 (1991), 264-269.
1534. PICKOVER, C. A. Computer-drawn faces characterizing nucleic acid sequences. *J. Mol. Graphics* 2 (1984), 107-110.
1535. PICKOVER, C. A. Frequency spectra of DNA sequences: application to a human bladder cancer gene. *J. Mol. Graphics* 2 (1984), 50.
1536. PICKOVER, C. A. DNA vectorgrams: Representation of cancer genes as movements on a 2D cellular lattice. *IBM J. Res. Develop.* 31(1) (1987), 111-119.
1537. PIERNO, G. N., BARNI, N., CANDURRO, M., CIPOLLARO, M., FRANZE, A., JULIANO, L., MACCHIATO, M. F., MASTROCINQUE, G., MOSCATELLI, C., SCARLATO, V., TRAMONTANO, A., AND CASCINO, A. Computer programs for the characterization of protein coding genes. *Nucl. Acids Res.* 12 (1984), 281-285.
1538. PIETROKOVSKI, S., HIRSHON, J., AND TRIFONOV, E. N. Linguistic measure of taxonomic and functional relatedness of nucleotide sequences. *J. Biomol. Struct. Dyn.* 7 (1990), 1251-1268.
1539. PINA, B., TRUSS, M., OHLENBUSCH, H., POSTMA, J., AND BEATO, M. DNA rotational positioning in a regulatory nucleosome is determined by base sequence. An algorithm to model the preferred superhelix. *Nucl. Acids Res.* 18(23) (1990), 6981-6987.
1540. PINES, M. Mapping the human genome, 1989. preprint.
1541. PIPAS, J. M., AND McMAHON, J. E. Method for predicting RNA secondary structure. *Proc. Natl. Acad. Sci. USA* 72 (1975), 2017-2021.
1542. PITTS, W. R., AND GOODFELLOW, J. M. Modelling of solvent positions around polar groups in proteins. *Protein Engineering* 4(5) (1991), 531-537.
1543. PIVEC, L., ROZKOT, F., SAZLOVA, P. S., AND VITEK, A. Consensus symmetry pattern in *E. coli* promoter sequences. *Folia Biologica* 31 (1985), 213-234.
1544. PIZZI, E., ATTIMONELLI, M., LIUNI, S., FRONTALI, C., AND SACCONI, C. A simple method for global sequence comparison. *Nucl. Acids. Res.* 20(1) (1992), 131-136.
1545. PIZZI, E., LIUNI, S., AND FRONTALI, C. Detection of latent sequence periodicities. *Nucl. Acids Res.* 18 (1990), 3745-3752.
1546. PLASKON, R. R., AND WARTELL, R. M. Sequence distributions associated with DNA curvature are found upstream of strong *E. coli* promoters. *Nucl. Acids Res.* 15 (1987), 785-796.
1547. PLATT, E., AND ROBSON, B. Calculation of biomolecular conformation, (in) *Computing in Biological Sciences*, Grisow, M.J., Barret, A.N., (eds.), (Elsevier, North Holland), pp.91-131, 1983.
1548. POHL, W. F. DNA and differential geometry. *The Math. Intel.* 3 (1980), 20-27.
1549. POLNER, G., DORGAI, L., AND OROSZ, L. PMAP, PMAPS: DNA physical map constructing programs. *Nucl. Acids Res.* 12(1) (1984), 227-236.
1550. POLOZOV, R. V., PANCHENKO, L. A., ESIPOVA, N. G., AND TUMANIAN, V. G. Classification of genomic regulatory sites given their primary and secondary structures, (in) *Human Genome-90*, Bayev, A.E., (ed.), (Abstracts of the 1st National Conference, Pereslavl-Zalesky), (Moscow), pp.208-209, 1990.
1551. POLOZOV, R. V., SMETANICH, Ya. S., AND SUKHORUKOV, B. I. Determination of the primary structure of linear heteropolymers. *Biofizika*. 17 (1972), 688-690.

1552. PONDER, J. W., AND RICHARDS, F. M. Tertiary templates for proteins: use of packing criteria in the enumeration of allowed sequences for different structural classes. *J. Mol. Biol.* **193** (1987), 775-791.
1553. PONGOR, S. The use of structural profiles and parametric sequence comparison in the rational design of polypeptides, (in) *Methods in Enzymology*, Vol.154, (Academic Press, NY), pp.459-473, 1987.
1554. PONGOR, S. Novel databases for molecular biology. *Nature* **323** (1988), 24.
1555. PONGOR, S., GUTTIERI, M. J., COHEN, L. M., AND SZALAY, A. A. Apple Pascal programs for prediction and comparative evaluation of protein secondary structure from nucleotide sequence data: application to ribulose-1, 5-bisphosphate carboxylase sequences. *DNA* **3** (1985), 319-326.
1556. PONGOR, S., GUTTIERI, M. J., COHEN, L. M., AND SZALAY, A. A. Microcomputer programs for prediction and comparative evaluation of protein secondary structure from nucleotide sequence data: application to ribulose-1, 5-bisphosphate carboxylase sequences. *DNA* **3** (1985), 319-326.
1557. PONGOR, S., AND SZALAY, A. A. Prediction of homology and divergence in the secondary structure of polypeptides. *Proc. Natl. Acad. Sci. USA* **82** (1985), 366-370.
1558. PONNUSWAMY, P. K., PRABHAKARAN, M., AND MANAVALAN, P. Hydrophobic packing and spatial arrangement of amino acid residues in globular proteins. *Biochim. Biophys. Acta* **623** (1980), 301-316.
1559. PONOMARENKO, M. P., AND ORLOV, Yu. L. Expert system of fast estimation of biopolymers characteristics utility for their classification, (in) *Computer Analysis of Structure, Function, and Evolution of Genetic Macromolecules. Intellectualization Problems*, Kolchanov, N.A., (ed.), (Novosibirsk), pp.197-220, 1989.
1560. PONOMARENKO, M. P., AND ORLOV, Yu. L. Expert systems for fast search of characteristics for biopolymers classification, (in) *Modelling and Computer Methods in Molecular Biology and Genetics*, Kolchanov, N.A., Ratner, V.A., (eds.), (Abstracts of the International Conference, Novosibirsk), pp.92-93, 1990.
1561. POPOWICZ, A. M., AND DASH, P. F. SIGSEQ: a computer program for predicting signal sequence cleavage sites. *CABIOS* **4** (1988), 405-406.
1562. PORTER, B., ACKER, L., LESTER, J., MURRAY, K., AND SOUTHER, A. The construction of a large-scale multifunctional knowledge base in botany, 1989. preprint.
1563. PORTER, T. Spherical shading. *Comp. Graphics* **12** (1978), 282.
1564. POSFAI, J., BHAGWAT, A. S., POSFAI, G., AND ROBERTS, R. J. Predictive motifs derived from cytosine methyltransferases. *Nucl. Acids Res.* **17**(7) (1989), 2421-2435.
1565. POWELL, P. A. Sequence similarity algorithms and a linear systolic VLSI implementation, 1989.
1566. POWELL, P. A. A further improved LCA algorithm, 1990. preprint.
1567. POWELL, P. A. RESIM—Fast algorithms for finding the similarity of regular expression based patterns and sequences. University of Minnesota Computer Science Department Technical Report #90-16, 1990.
1568. POWELL, P. A. Serial and parallel algorithms for rapid sequence similarity determination, 1990. preprint.
1569. PRAGER, E. M., AND WILSON, A. C. Construction of phylogenetic trees for proteins and nucleic acids: empirical evaluation of alternative matrix methods. *J. Mol. Evol.* **11** (1978), 129-142.
1570. PRAMANIK, S., AND KING, C. T. A hardware pattern matching algorithm on a dataflow. *The Computer Journal* **38** (1985), 264-269.
1571. PRESNELL, S. R., AND BENNER, S. A. The design of synthetic genes. *Nucl. Acids Res.* **16** (1988), 1693-1702.
1572. PRESSING, J., AND REANNY, D. C. Divided genomes and intrinsic noise. *J. Mol. Evol.* **20** (1984), 135-146.
1573. PRESTA, L. G. Protein structure analysis and development of databases. *Protein Engineering* **2** (1989), 395-397.
1574. PRESTA, L. G., AND ROSE, D. G. Helix signals in proteins. *Science* **240** (1988), 1632-1641.
1575. Proceedings of the Workshop on Advanced Computer Technologies and Biological Sequencing, ANL-88-45, November, 3-5, 1988.
1576. Protein Engineering Club Database Group. A Protein sequence/structure database, *Nature*, **335**:745-746, 1988.
1577. PTITSYN, O. B. Protein folding: general physical model. *FEBS Lett.* **131** (1981), 197-202.
1578. PTITSYN, O. B., AND FINKELSTEIN, A. V. Theory of protein secondary structure and algorithm of its prediction. *Biopolymers* **22** (1983), 15-25.
1579. PUSTELL, J., AND KAFATOS, F. C. A high speed, high capacity homology matrix: Zooming through SV40 and polyoma. *Nucl. Acids Res.* **10** (1982), 4765-4782.
1580. PUSTELL, J., AND KAFATOS, F. C. A convenient and adaptable package of computer programs for DNA and protein sequence management, analysis, and homology determination. *Nucl. Acids Res.* **12** (1984), 643-655.
1581. PUSTELL, J. M. Interactive molecular biology computing. *Nucl. Acids Res.* **16** (1988), 1813-1820.
1582. QIAN, N., AND SEJNOWSKI, T. J. Predicting the secondary structure of globular proteins using neural network models. *J. Mol. Biol.* **202** (1988), 865-884.
1583. QUEEN, C., AND KORN, L. G. Computer analysis of nucleic acids and proteins, (in) *Methods in Enzymology*, Vol.65, (Academic Press, NY), pp.595-609, 1980.

1584. QUEEN, C., AND KORN, L. J. A comprehensive sequence analysis program for the IBM personal computer. *Nucl. Acids Res.* **12**(1) (1984), 581-599.
1585. QUEEN, C., WEGMAN, M. N., AND KORN, L. J. Improvements to a program for DNA analysis: A procedure to find homologies among many sequences. *Nucl. Acids Res.* **10**(1) (1982), 449-457.
1586. QUIGLEY, G. J., GEHRKE, L., ROTH, D. A., AND AURON, P. E. Computer-aided nucleic acid secondary structure modelling incorporating enzymatic digestion data. *Nucl. Acids Res.* **12**(1) (1984), 347-366.
1587. QUINQUETON, J., AND MOREAU, J. Application of learning techniques to splicing site recognition. *Biochimie* **67** (1985), 541-547.
1588. RACKOWSKY, S., AND GOLDSTEIN, D. A. Protein comparison and classification: a differential geometrical approach. *Proc. Natl. Acad. Sci. USA* **85** (1988), 777-781.
1589. RADDING, C. M. Genetic recombination strand transfer and mismatch repair. *Ann. Rev. Biochem.* **47** (1978), 847.
1590. RAGONE, R., FACCHIANO, F., FACCHIANO, A., FACCHIANO, A. M., AND COLONNA, G. Flexibility plot of proteins. *Protein Engineering* **2** (1989), 497-504.
1591. RAIHA, K. J., AND UKKONEN, E. The shortest common supersequence problem over binary alphabet is NP-complete. *Theoretical Computer Science* **16** (1990), 187-198.
1592. RAIHA, L. Approximate sequence comparison: A study with histograms. *Pattern Recognition* **12** (1990), 159-169.
1593. RALPH, W. W., WEBSTER, T., AND SMITH, T. F. A modified Chou and Fasman protein structure algorithm. *CABIOS* **3**(3) (1987), 211-216.
1594. RANDIC, M. On molecular identification numbers. *J. Chem. Inf. Comput. Sci.* **24** (1984), 164-175.
1595. RANDIC, M., BRISSEY, G. M., SPENCER, R. B., AND WILKINS, C. L. Searchfor all self-avoiding paths for molecular graphs. *Computers and Chemistry* **3** (1979), 5-13.
1596. RANDIC, M., AND WILKINS, C. L. Graph theoretical approach to recognition of structural similarity in molecules. *J. Chem. Inf. Comput. Sci.* **19** (1979), 31-37.
1597. RANDIC, M., AND WILKINS, C. L. On a graph theoretical basis for ordering of structures. *Chemical Physics Letters* **63**(2) (1979), 332-336.
1598. RANKA, S., AND HEYWOOD, T. Two-dimensional pattern matching with k mismatches. *Pattern Recognition* **24** (1991), 31-40.
1599. RAO, G. S., HAMID, Z., AND RAO, J. S. The information content of DNA and evolution. *J. Theor. Biol.* **81** (1979), 803-807.
1600. RAO, J. K. M. New scoring matrix for amino acid residue exchanges based on residue characteristic physical parameters. *Int. J. Pept. Protein Res.* **29** (1987), 276-281.
1601. RAO, J. K. M., AND ARGOS, P. A conformational preference parameter to predict helices in integral membrane proteins. *Biochimica et Biophysica Acta* **869** (1986), 197-214.
1602. RATNER, V. A. Genetic language and adjacent problems, (in) **Modelling and Computer Methods in Molecular Biology and Genetics**, Kochanov, N.A., Ratner, V.A., (eds.), (Abstracts of the International Conference, Novosibirsk), pp.156-157, 1990.
1603. RAUPACH, R. E. Computer programs used to aid in the selection of DNA hybridization probes. *Nucl. Acids Res.* **12** (1984), 833-836.
1604. RAWLING, N., ASHMAN, K., AND WITTMANN-LIEBOLD, B. Computerized version of Chou and Fasman protein secondary structure predictive method. *Int. J. Pept. Protein Res.* **22** (1983), 515-524.
1605. RAWLINGS, C. J. **Software Directory for Molecular Biologists**, Macmillan Publishers, NY, 1986.
1606. RAWLINGS, C. J. Designing databases for molecular biology. *Nature* **334** (1988), 477.
1607. RAWLINGS, C. J. Databases, artificial intelligence and knowledge-based systems for molecular biology. *Biochem. Soc. Trans.* **17** (1989), 851-855.
1608. RAWLINGS, C. J., TAYLOR, W. R., NYAKAIRU, J., FOX, J., AND STERNBERG, M. J. E. Reasoning about protein topology using the logic programming language PROLOG. *J. Mol. Graphics* **3** (1985), 151-157.
1609. RAWLINS, D. J., AND SHAW, P. J. Localization of ribosomal and telomeric DNA sequences in intact plant nuclei by *in-situ* hybridization and three-dimensional optical microscopy. *J. Microscopy* **157** (1990), 83-89.
1610. RECHID, R., VINGRON, M., AND ARGOS, P. A new interactive protein sequence alignment program and comparison of its results with widely used algorithms. *CABIOS* **5**(2) (1989), 107-113.
1611. REDDADAY, S. F., AND PAGE, R. M. R. High speed data searching with a processor array. *Microprocessing and Microprogramming* **24** (1988), 655-660.
1612. REEKE, Jr., George, N. Protein folding: Computational approaches to an exponential-time problem. *Ann. Rev. Comput. Sci.* **3** (1988), 59-84.
1613. REICH, J. G., AND MEISKE, W. A simple statistical significance test of window scores in large dot matrices obtained from protein or nucleic acid sequences. *CABIOS* **3** (1987), 25-30.
1614. REICHERT, T. A., COHEN, D. N., AND WONG, A. K. C. An application of information theory to genetic mutations and matching of polypeptide sequences. *J. Theor. Biol.* **42** (1973), 245-261.
1615. REIF, J. H. An optimal parallel algorithm for integer sorting, (in) **IEEE Symposium on the Foundations of Computer Science**, pp.496-504, 1985.

1616. REISNER, A. H., AND BUCHOLTZ, C. A. Utilization of sequence libraries on a 16 bit mini-computer with a particular reference to high speed searching. *Nucl. Acids Res.* **12** (1984), 409-416.
1617. REISNER, A. H., AND BUCHOLTZ, C. A. Two-dimensional abstract representations (signatures) of proteins. *CABIOS* **3** (1987), 171-175.
1618. REISNER, A. H., AND BUCHOLTZ, C. A. The use of various properties of amino acids in color and monochrome dot-matrix analyses for protein homologies. *CABIOS* **3** (1988), 395-402.
1619. REISNER, A. H., AND WESTWOOD, N. H. Assessment of similarities of pairs and groups of proteins using transformed amino acid-residue data. *J. Mol. Evol.* **18** (1982), 240-250.
1620. REMINGTON, S. J., AND MATTHEWS, B. W. A systematic approach to the comparison of protein structure. *J. Mol. Biol.* **140** (1980), 77-99.
1621. RENAN, M. J. Conserved 12-bp element downstream from mRNA polyadenylation sites. *Gene* **60** (1987), 245-254.
1622. REYES, V. E., LEW, R. A., LU, S., AND HUMPHREYS, R. E. Prediction of α helices and T cell-presented sequences in proteins with algorithms based on strip-of-helix hydrophobicity index, (in) *Methods in Enzymology-Molecular design and Modeling: Concepts and Applications*, Part A, Vol.202, Langone, J.J., (ed.), (Academic Press, NY), pp.225-238, 1991.
1623. REYES, V. E., PHILLIPS, L., HUMPHREYS, R. E., AND LEW, R. A. Prediction of protein helices with a derivative of the strip-of-helix hydrophobicity algorithm. *J. Biol. Chemistry* **264**(22) (1989), 12854-12858.
1624. RHODES, D., AND KLUG, A. Sequence-dependent helical periodicity of DNA. *Nature* **292** (1981), 378.
1625. RHODES, D., AND KLUG, A. An underlying repeat in some transcriptional control sequences corresponding to half a double helical turn of DNA. *Cell* **46** (1986), 123-132.
1626. RICHARDS, F. M., AND KUNDROT, C. E. Identification of structural motifs from protein coordinate data: Secondary structure and first-level supersecondary structure. *PROTEINS: Structure, Function, and Genetics* **3** (1988), 71-84.
1627. RICHARDSON, J. S. The anatomy and taxonomy of protein structure. *Adv. Prot. Chem.* **34** (1981), 167-339.
1628. RICHARDSON, J. S. Describing patterns of protein tertiary structure, (in) *Methods in Enzymology*, Vol.115, (Academic Press, NY), pp.341-358, 1985.
1629. RICHARDSON, J. S. Schematic drawing of protein structures, (in) *Methods in Enzymology*, Vol.115, (Academic Press, NY), pp.359-380, 1985.
1630. RICHARDSON, J. S., AND RICHARDSON, D. C. Amino acid preference for specific locations at the end of alpha-helices. *Science* **240** (1988), 1648-1653.
1631. RINSMA, I., HENDY, M., AND PENNY, D. Distribution of the number of matches between nucleotide sequences. *Bull. Math. Biol.* **52** (1990), 349-358.
1632. RIPPmann, F., AND TAYLOR, W. R. Visualization of structural similarity in proteins. *J. Mol. Graphics* **9**(3) (1991), 169-174.
1633. RISLER, J. L., DELORME, M. O., DELACROIX, H., AND HENAUT, A. Amino acid substitutions in structurally related proteins. A pattern recognition approach. *J. Mol. Biol.* **240** (1988), 1019-1029.
1634. ROBBINS, R. J. Challenges in the human genome project. *IEEE Eng. Med. Biol.* **11**(1) (1992), 25-34.
1635. ROBERTS, L. New chip may speed genome analysis. *Science* **244** (1989), 655-656.
1636. ROBERTS, L. The worm project. *Science* **248** (1989), 655-656.
1637. ROBSON, B. Analysis of the code relating sequence to conformation in globular proteins - theory and application of expected information. *Biochem. J.* **141** (1974), 853-867.
1638. ROBSON, B., DOUGLAS, G. M., AND GARNIER, J. A simple method and BASIC program for secondary structure prediction on a microcomputer, (in) *Computing in Biological Sciences*, Grisow, M.J., Barret, A.N., (eds.), (Elsevier, North Holland), pp.132-142, 1983.
1639. ROBSON, B., AND OSGUTHORPE, D. J. Refined models for computer simulation of protein folding. *J. Mol. Biol.* **132** (1979), 19-51.
1640. ROBSON, B., PLATT, E., FINN, P., MILLARD, P., GIBRAT, J.-F., AND GARNIER, J. Prediction of the conformation and antigenic determinants of the V-sis viral oncogene product homologous with human platelet-derived growth factor. *Int. J. Pept. Protein Res.* **25** (1985), 1-8.
1641. RODIER, F., GABARRO-ARPA, J., EHRLICH, R., AND REISS, C. Key for protein coding sequence identification: Computer analysis of codon strategy. *Nucl. Acids Res.* **10** (1982), 391-402.
1642. RODIER, F., AND SALLANTIN, J. Localization of the initiation of translation in messenger RNAs of prokaryotes by learning techniques. *Biochimie* **67** (1985), 533-539.
1643. ROE, B. R. Computer programs for molecular biology: an overview of DNA sequencing and protein analysis packages. *Biotechniques* **6** (1988), 560-563.
1644. ROGERS, S., WELLS, R., AND RECHSTEINER, M. Amino acid sequences common to rapidly degraded proteins: the PEST hypothesis. *Science* **234** (1986), 364-368.
1645. ROGOZIN, I. B., SOLOVYOV, V. V., AND KOLCHANOV, N. A. Method of analysis of oligonucleotide connection with recombinational events, (in) *Modelling and Computer Methods in Molecular Biology and Genetics*, Kochanov, N.A., Ratner, V.A., (eds.), (Abstracts of the International Conference, Novosibirsk), pp.49-50, 1990.
1646. ROHDE, K. A multiple alignment routine for proteins: use, limits and experiences, (in) *Abstracts of Application of Data Bases and Microcomputers in Molecular Biology*, (Jena), p.47, 1988.

1647. ROTBERG, M. A. Search for similar fragments in several sequences, (in) *Human Genome-90*, Bayev, A.E., (ed.), (Abstracts of the 1st National Conference, Pereslavl-Zalesky, Moscow), pp.209-210, 1990.
1648. ROIZES, G., AND PELAQUIER, E. A method for detecting the coding DNA for a protein of known sequence in a collection of chimeric plasmids constructed with cDNAs. *Nucl. Acids Res.* 8(2) (1980), 351-360.
1649. ROMERO-HERRERA, A. E., LIESKA, N., GOODMAN, M., AND SIMMONS, E. L. The use of amino acid sequence analysis in assessing evolution. *Biochimie* 61 (1979), 767-779.
1650. ROODE, D., LIEBSCHUTZ, R., MAULIK, S., AND FRIEDEMANN, T. New developments at BIONET. *Nucl. Acids Res.* 16(5) (1988), 1857-1859.
1651. ROOMAN, M. J., RODRIGUEZ, J., AND WODAK, S. J. Automatic definition of recurrent local structure motifs in proteins. *J. Mol. Biol.* 213(2) (1990), 327-336.
1652. ROOMAN, M. J., AND WODAK, S. Identification of predictive sequence motifs limited by protein structure data base size. *Nature* 335 (1988), 45-49.
1653. ROSE, D., AND Langley, P. Chemical discovery as belief revision. *Machine Learning* 1 (1986), 423-451.
1654. ROSE, G. D. Prediction of chain turns in globular proteins on a hydrophobic basis. *Nature* 272 (1978), 586-590.
1655. ROSE, G. D., GESELOWITZ, A. R., LESSER, G. J., LEE, R. H., AND ZEHFUS, M. H. Hydrophobicity of amino acid residues in globular proteins. *Science* 229 (1985), 834-838.
1656. ROSE, G. D., GIERASCH, L. M., AND SMITH, J. A. Turns in peptides and proteins. *Adv. Prot. Chem.* 387 (1985), 1-109.
1657. ROSE, G. D., AND ROY, S. Hydrophobic basis of packing in globular proteins. *Proc. Natl. Acad. Sci. USA* 77 (1980), 4643-4647.
1658. ROSE, G. D., WINTER, R. H., AND WETLAUFER, D. B. A testable model for protein folding. *FEBS Lett.* 63 (1976), 10-16.
1659. ROSEN, G. Rook's tour representation of the genetic code. *Bull. Math. Biol.* 53(6) (1991), 845-851.
1660. ROSEN, K. H. *Elementary Number Theory and Its Applications*, Addison-Wesley, Reading, MA, 1984.
1661. ROSEN, R. Pattern generation in networks, (in) *Progress in Theoretical Biology*, Vol. 6, Rosen, R., (ed.), (Academic Press, NY), 1981.
1662. ROSS, A. M., AND GOLUB, E. E. A computer graphics program system for protein structure representation. *Nucl. Acids Res.* 16 (1988), 1801-1812.
1663. ROSSMANN, M. G., AND ARGOS, P. Exploring structural homology of proteins. *J. Mol. Biol.* 105 (1976), 75-95.
1664. ROSSMANN, M. G., AND ARGOS, P. Three-dimensional coordinates from stereo diagrams of molecular structures. *Acta Crystallogr.* 36 (1980), 819-823.
1665. ROSSMANN, M. G., AND ARGOS, P. Protein folding. *Ann. Rev. Biochem.* 50 (1981), 497-532.
1666. ROWE, G. W. A three-dimensional representation for base composition of protein-coding DNA sequences. *J. Theor. Biol.* 112 (1985), 433-444.
1667. ROWE, G. W., AND TRAINOR, L. E. H. On the informational content of viral DNA. *J. Theor. Biol.* 101 (1983), 151-170.
1668. ROWE, G. W., AND TRAINOR, L. E. H. A thermodynamic theory of codon bias in viral genes. *J. Theor. Biol.* 101 (1983), 171-203.
1669. ROWLETT, R. A low cost system for the graphical display of space-filling models of proteins. *Computers and Chemistry* 9 (1985), 301-315.
1670. ROY, A., AND ROY, T. K. A database management system for recombinant DNA clones and hosts for use in IBM personal computers. *CABIOS* 5 (1989), 321-322.
1671. ROYTBERG, M. A. A search for common patterns in many sequences. *CABIOS* 8(1) (1992), 57-64.
1672. ROZKOT, F., SAZELOVA, P., AND PIVEC, I. A novel method for promoter search enhanced by function-specific subgrouping of promoters-developed and tested on *E. coli* system. *Nucl. Acids Res.* 17 (1989), 4799-4815.
1673. RUDD, K., MILLER, W., WERNER, C., OSTELL, J., TOLSTOSHEV, C., AND SATTERFIELD, S. G. Mapping sequenced *E. coli* genes by computer: software, strategies and examples. *Nucl. Acids Res.* 19(3) (1991), 637-647.
1674. RUDD, K. E., MILLER, W., OSTELL, J., AND BENSON, D. A. Alignment of *Escherichia coli* K-12 DNA sequences to a genomic restriction map. *Nucl. Acids Res.* 18, 2 (1990), 313-321.
1675. RUSSEK-COHEN, E., AND JACOBS, D. Detecting outliers in a 5S rRNA database. *Binary Comput. Microbiol.* 1 (1989), 115-123.
1676. RUSSELL, G. C. DNA SIZE-a program to calculate the sizes of DNA fragments in agarose gels from the mobility of standard fragments. *Binary* 3 (1984), 31-32.
1677. RUSSELL, P. J., CRANDALL, R. E., AND FEINBAUM, R. GELSIS: Pascal-implemented analysis of one-dimensional electrophoresis gels. *Nucl. Acids Res.* 12(1) (1984), 493-498.
1678. RYTTER, W., AND DIKS, K. On optimal parallel computations for sequences of brackets, (in) *Sequences, Combinatorics, Compression, Security and Transmission*, Capocelli, R.M., (Springer-Verlag), pp.92-105, 1990.

1679. SACCOONE, C., LANAVE, C., PESOLE, G., AND PREPARATA, G. Influence of base composition on quantitative estimates of gene evolution, (in) *Methods in Enzymology-Molecular Evolution: Computer Analysis of Protein and Nucleic Acid Sequences*, Vol.183, Doolittle, R.F., (ed.), (Academic Press, NY), pp.570-583, 1990.
1680. SACKIN, M. J. Cross-association: A method of comparing protein sequences. *Biochem. Genetics* 5 (1971), 287-313.
1681. SADLER, J. R., WATERMAN, M. S., AND SMITH, T. F. Regulatory pattern identification in nucleic acid sequences. *Nucl. Acids Res.* 11 (1983), 2221-2231.
1682. SAITOU, N. Maximum likelihood methods, (in) *Methods in Enzymology-Molecular Evolution: Computer Analysis of Protein and Nucleic Acid Sequences*, Vol.183, Doolittle, R.F., (ed.), (Academic Press, NY), pp.584-598, 1990.
1683. SAITOU, N., AND NEI, M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* 4 (1987), 406-425.
1684. SALEMME, A., AND FURANO, A. V. A convenient method for locating sets of related short sequences in DNA sequences of any length. *Nucl. Acids Res.* 12 (1984), 257-262.
1685. SALEMME, F. R. A model for catabolite activator protein binding to supercoiled DNA. *Proc. Natl. Acad. Sci.* 79 (1982), 5263.
1686. SALINAS, J., MATASSI, G., MONTERO, L., AND BERNARDI, G. Compositional compartmentalization and compositional patterns in the nuclear genomes of plants. *Nucl. Acids Res.* 16 (1988), 4269-4285.
1687. SALINAS, J., ZERIAL, M., FILIPSKI, J., AND BERNARDI, G. Gene distribution and nucleotide sequence organization in the mouse genome. *Eur. J. Biochem.* 160 (1986), 469-478.
1688. SALLANTIN, J., HAIECH, J., AND RODIER, F. Search for promoter sites of prokaryotic DNA using learning techniques location. *Biochimie* 67 (1985), 549-553.
1689. SALLANTIN, J., HAIECH, J., AND RODIER, F. Search for promoter sites of prokaryotic DNA using learning techniques. *Biochimie* 47 (1985), 549-553.
1690. SANDA, A. I., AND FORD, J. P. Genomic analysis I: Inheritance units and genetic selection in the rapid discovery of locus linked DNA markers. *Nucl. Acids Res.* 14(80) (1986), 7265-7283.
1691. SANDERS, J. Z., PETTERSON, A. A., HUGHES, P. J., CONNELL, C. R., RAFF, M., MENCHEN, S., HOOD, L. E., AND TEPLow, D. B. Imaging as a tool for improving length and accuracy of sequence analysis in automated fluorescence-based DNA sequencing. *Electrophoresis* 12 (1991), 3-11.
1692. SANGER, F., COULSON, A. R., BARRELL, B. G., SMITH, A. J. H., AND ROE, B. A. Cloning in single-stranded bacteriophage as an aid to rapid DNA sequencing. *J. Mol. Biol.* (1980), 161-178.
1693. SANGER, F., NICKLEN, S., AND COULSON, A. R. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci.* 74(12) (1977), 5463-5467.
1694. SANKOFF, D. Matching sequences under deletion/insertion constraints. *Proc. Natl. Acad. Sci. USA* 69 (1972), 4-6.
1695. SANKOFF, D. Minimal mutation trees of sequences. *SIAM J. Appl. Math.* 28 (1975), 35-42.
1696. SANKOFF, D. Simultaneous solution of the RNA folding, alignment and protosequence problems. *SIAM J. Appl. Math.* 45 (1985), 810-824.
1697. SANKOFF, D., CEDERGREN, R., AND ABEL, Y. Genomic divergence through gene rearrangement, (in) *Methods in Enzymology-Molecular Evolution: Computer Analysis of Protein and Nucleic Acid Sequences*, Vol.183, Doolittle, R.F., (ed.), (Academic Press, NY), pp.428-438, 1990.
1698. SANKOFF, D., AND CEDERGREN, R. J. Simultaneous comparison of three or more sequences related by a tree, (in) *Time Warps, string Edits, and Macromolecules: The Theory and Practice of Sequence Comparison*, Sankoff, D., Kruskal, J.B., (eds.), (Addison-Wesley, Reading), pp.253-263, 1983.
1699. SANKOFF, D., CEDERGREN, R. J., AND MCKAY, W. A strategy for sequence phylogeny research. *Nucl. Acids Res.* 10(1) (1982), 421-431.
1700. SANKOFF, D., AND GOLDSTEIN, M. Probabilistic models of genome shuffling. *Bull. Math. Biol.* 51(1) (1989), 117-124.
1701. SANKOFF, D., KRUSKAL, J. B., MAINVILLE, S., AND CEDERGREN, R. J. Fast algorithms to determine RNA secondary structures containing multiple loops, (in) *Time Warps, String Edit, and Macromolecules: The Theory and Practice of Sequence Comparison*, Sankoff, D., Kruskal, J.B., (eds.), (Addison-Wesley, Reading, Massachusetts), 1983.
1702. SANKOFF, D., AND SELLER, P. Shortcuts, diversions, and maximal chains in partially ordered sets. *Discrete Math.* 4 (1973), 287-293.
1703. SANTIBANEZ, M., AND ROHDE, K. A multiple alignment program for protein sequences. *CABIOS* 3 (1987), 111-114.
1704. SARAI, A., MAZUR, J., NUSSINOV, R., AND JERNIGAN, R. L. Origin of DNA helical structure and its sequence dependence. *Biochemistry* 27 (1988), 8498-8502.
1705. SARAI, A., MAZUR, J., NUSSINOV, R., AND JERNIGAN, R. L. Sequence dependence of DNA conformational flexibility. *Biochemistry* 28 (1989), 7842-7849.
1706. SARGAN, D. R., GREGORY, S. P., AND BUTTERWORTH, P. H. W. A possible novel interaction between the 3'-end of 18S ribosomal RNA and the 5'-leader of many eukaryotic messenger RNAs. *FEBS Lett.* 147 (1982), 133-136.

1707. SAROFF, H. The uniqueness of protein sequences: Uniqueness diagrams for the Dayhoff file 1984. *Bull. Math. Biol.* **46** (1984), 661–672.
1708. SAROFF, H. A. A note on the evaluation of similarity (homology) of short sequences with long sequences. *Bull. Math. Biol.* **46** (1984), 951–961.
1709. SASAGAWA, T., OKUYAMA, T., AND TELLER, D. C. Prediction of peptide retention times in reversed-phase high-performance liquid chromatography during linear gradient elution. *J. Chromatography* **240** (1982), 329–340.
1710. SATCHWELL, S. C., DREW, H. R., AND TRAVERS, A. A. Sequence periodicities in chicken nucleosome core DNA. *J. Mol. Biol.* **191** (1986), 659–675.
1711. SAURIN, W. Repetitive palindromic sequences in *Escherichia coli*. Detection and characterization with a new computer program. *CABIOS* **3** (1987), 121–127.
1712. SAURIN, W., AND MARLIERE, P. Matching relational patterns in nucleic acid sequences. *CABIOS* **3** (1987), 115–120.
1713. SCHABACK, R. On the expected sublinearity of the Boyer-Moore algorithm. *SIAM J. Computing* **17** (1988), 648–658.
1714. SCHAEFFER, E., AND SNINSKY, J. J. Predicted secondary structure similarity in the absence of primary amino acid sequence homology: hepatitis B virus open reading frames. *Proc. Natl. Acad. Sci. USA* **81** (1984), 2902–2906.
1715. SCHAFFER, H. E., AND SEDEROFF, R. R. Improved estimation of DNA fragment lengths from agarose gels. *Anal. Biochem.* **115** (1981), 113–122.
1716. SCHERAGA, H. A. Recent progress in the theoretical treatment of protein folding. *Biopolymers* **22** (1983), 1–14.
1717. SCHERER, G. E. F., WALKINSHAW, M. D., AND ARNOTT, S. A computer aided oligonucleotide analysis provides a model sequence for RNA polymerase-promoter recognition in *E. coli*. *Nucl. Acids Res.* **5** (1978), 3759.
1718. SCHERER, G. F. E., WALKINSHAW, M. D., ARNOTT, S., AND MORRE, D. J. The ribosome binding sites recognized by *E. coli* ribosomes have regions with signal character in both the leader and protein coding segments. *Nucl. Acids Res.* **8** (1980), 3895–3907.
1719. SCHIEBER. On finding lowest common ancestors: Simplification and parallelization. *SIAM J. Computing* **17** (1988), 1253–1262.
1720. SCHIFFER, M., AND EDMUNSON, A. B. Use of helical wheels to represent the structures of proteins and to identify segments with helical potential. *Biophys. J.* **7** (1967), 121–135.
1721. SCHLICK, T., AND FIGUEROA, S. A molecular dynamics simulation of a water droplet by the implicit-Euler/Langevin scheme. *J. Chem. Phys.* **94**(3) (1991), 2118–2129.
1722. SCHLICK, T., AND OVERTON, M. A Powerful Truncated Newton Method for Potential Energy Minimization. *J. Computational Chemistry* **8**(7) (1987), 1025–1039.
1723. SCHMATCHENKO, V. V., AND BEREZHNEV, A. B. Mapping of functional regions by graphic representation of long DNA sequences, (in) *Modelling and Computer Methods in Molecular Biology and Genetics*, Kolchanov, N.A., Ratner, V.A., (eds.), (Abstracts of the International Conference, Novosibirsk), p.13, 1990.
1724. SCHMIDT, M. C., KAO, C., PEI, R., AND BERK, A. J. Yeast TATA-box transcription factor gene. *Proc. Natl. Acad. Sci.* **86** (1989), 7785–7789.
1725. SCHNEIDER, T. D., AND STEPHENS, R. M. Sequences logos: a new way to display consensus sequences. *Nucl. Acids Res.* **17** (1990), 6097–6100.
1726. SCHNEIDER, T. D., AND STORMO, G. D. Excess information at bacteriophage T7 genomic promoters detected by a random cloning technique. *Nucl. Acids Res.* **17** (1989), 659–674.
1727. SCHNEIDER, T. D., STORMO, G. D., GOLD, L., AND EHRENFEUCHT, A. Information content of binding sites in nucleotide sequences. *J. Mol. Biol.* **188** (1986), 415–431.
1728. SCHNEIDER, T. D., STORMO, G. D., HAEMER, J. S., AND GOLD, L. A design for computer nucleic acid sequence storage, retrieval, and manipulation. *Nucl. Acids Res.* **10**(9) (1982), 3013–3024.
1729. SCHONHERR, G., AND NOOLANDI, J. Fluctuating bond model of DNA gel electrophoresis. *Electrophoresis* **12** (1991), 432–435.
1730. SCHONHERR, G., AND NOOLANDI, J. Fluctuating boud model of DNA gel electrophoresis. *Electrophoresis* **12** (1991), 432–435.
1731. SCHROEDER, J. L., AND BLATTNER, F. R. Least-squares methods for restriction mapping. *Gene* **4** (1978), 168–174.
1732. SCHROEDER, J. L., AND BLATTNER, F. R. Formal description of DNA oriented computer language. *Nucl. Acids Res.* **10**(1) (1982), 69–84.
1733. SCHULER, G. D., ALTSCHUL, S. F., AND LIPMAN, D. J. A workbench for multiple alignment construction and analysis. *Protein Struct. Func. Genet* **9**(3) (1991), 180–190.
1734. SCHULMEISTER, T., TSCHAPEK, A., AND ZINKE, M. Visualization of the entire surface of a protein by cartographic projection. *CABIOS* **2** (1986), 265–268.
1735. SCHULZ, G. E. A critical evaluation of methods for prediction of protein secondary structures. *Ann. Rev. Biophys. Biophys. Chem.* **17** (1988), 1–21.

1736. SCHULZ, G. E., AND SCHIRMER, R. H. Models, display, and documentation of protein structure, (in) **Principles of Protein Structure**, (Springer-Verlag, Berlin), pp.131-148, 1979.
1737. SCHULZ, G. E., AND SCHIRMER, R. H. Prediction of secondary structure from the amino acid sequence, (in) **Principles of Protein Structure**, (Springer-Verlag, Berlin), pp.108-130, 1979.
1738. SCHULZ, G. E., AND SCHIRMER, R. H. Protein evolution, (in) **Principles of Protein Structure**, (Springer-Verlag, Berlin), pp.166-205, 1979.
1739. SCHUTZ, H. Prediction of protein secondary structures by pattern recognition: the program COHEN, (in) **Abstracts of: Application of Data Bases and Microcomputers in Molecular Biology**, (Jena), p.48, 1988.
1740. SCHWARTZ, R. M., AND DAYHOFF, M. O. Matrices for detecting distant relationships, (in) **Atlas of Protein Sequence and Structure**, 5 suppl.3:353-358, (Nat. Biomed. Res. Found., Washington D.C.), 1978.
1741. SCHWINDINGER, W. F., AND WARNER, R. J. DNA sequence analysis on the IBM-PC. *Nucl. Acids Res.* **12** (1984), 601-604.
1742. SCOTT, D. W. Some thoughts on estimation and visualization of multi-dimensional surfaces. August 18, 1989. (Overheads only).
1743. SEARLS, D. B. Representing genetic information with formal grammars, (in) **Proceedings of the AAAI** **88**, pp.386-391, 1988.
1744. SEDGEWICK, R. *Algorithms*. Addison-Wesley, Reading, MA, 1983.
1745. SEGE, R. D., AND SAXBERG, B. E. H. A statistical test for comparing several nucleotide sequences. *Nucl. Acids Res.* **10**(1) (1982), 375-389.
1746. SEGE, R. D., SOLL, D., RUDDLE, F., AND QUEEN, C. L. A conversational system for the computer analysis of nucleic acid sequences. *Nucl. Acids Res.* **9** (1981), 437-444.
1747. SEJNOWSKI, T. J., AND ROSENBERG, C. R. *Complex Systems* **1** 145, 1987.
1748. SEL'KOV, E. E., GORYANIN, I. I., KAIMACHNIKOV, N. P., SHEVELEV, E. L., AND YUNUS, I. Data and knowledge banks on enzymes and metabolic pathways, (in) **Abstracts of: Application of Data Bases and Microcomputers in Molecular Biology**, (Jena), p.19, 1988.
1749. SELLERS, P. An algorithm for the distance between two finite sequences. *J. Combinatorial Theor. (A)* **16** (1974), 253-258.
1750. SELLERS, P. On the theory and computation of evolutionary distances. *SIAM J. Appl. Math.* **26** (1974), 787-793.
1751. SELLERS, P. H. Pattern recognition in genetic sequences. *Proc. Natl. Acad. Sci.* **76**(7) (1979), 3041.
1752. SELLERS, P. H. The theory and computation of evolutionary distances: pattern recognition. *J. Algorithms* **1** (1980), 359-373.
1753. SELLERS, P. H. Pattern recognition in genetic sequences by mismatch density. *Bull. Math. Biol.* **46**(4) (1984), 501-514.
1754. SENAPATHY, P. Origin of eukaryotic introns: A hypothesis, based on codon distribution statistics in genes, and its implications. *Proc. Natl. Acad. Sci. USA* **83** (1986), 2133-2137.
1755. SENAPATHY, P. Possible evolution of splice-junction signals in eukaryotic genes from stop codons. *Proc. Natl. Acad. Sci. USA* **85** (1988), 1129-1133.
1756. SENAPATHY, P., SHAPIRO, M. B., AND HARRIS, N. L. Splice junctions, branch point sites, and exons - sequence statistics, identification, and applications to Genome project, (in) **Methods in Enzymology-Molecular Evolution: Computer Analysis of Protein and Nucleic Acid Sequences**, Doolittle, R.F., (ed.), Vol.183, pp.252-281, 1990.
1757. SENAPATHY, P., SHAPIRO, M. B., AND HARRIS, N. L. Splice junctions, branch point sites, and exons: sequence statistics, identification, and applications to genome project, (in) **Methods in Enzymology-Molecular Evolution: Computer Analysis of Protein and Nucleic Acid Sequences**, Vol.183, Doolittle, R.F., (ed.), (Academic Press, NY), pp.252-278, 1990.
1758. SERESSIOTIS, A., AND BAILEY, J. E. MPS: An artificially intelligent software system for the analysis and synthesis of metabolic pathways. *Biotechnology and Bioengineering* **31** (1988), 587-602.
1759. SETO, Y., IHARA, S., KOHTSUKI, S., OOI, T., AND SAKAKIBARA, S. Peptide and protein databanks in Japan, (in) **Computational Molecular Biology. Sources and Methods for Sequence Analysis**, Lesk, A.M., (ed.), (Oxford University Press, Oxford), pp.27-37, 1988.
1760. SETO, Y., NOZAKI, Y., AND SAKAKIBARA, S. Data base for information of peptide (DBIP): efficient production of protein sequence data base and its problem. *Protein, Nucleic Acid and Enzyme* **29** (1986), 195-209.
1761. SHABALINA, S. A., KONDRAZOV, A. S., AND VERNOSLOV, S. E. Some characteristic features of 5'-regulatory regions of heat shock protein genes, (in) **Modelling and Computer Methods in Molecular Biology and Genetics**, Kolchanov, N.A., Ratner, V.A., (eds.), (Abstracts of the International Conference, Novosibirsk), pp.16-17, 1990.
1762. SHABALINA, S. A., YURIEVA, O. V., SPIRIDONOV, N. A., AND KONDRAZOV, A. S. Comparative analysis of textually similar sites in DNA regulatory regions, Preprint, Scientific Computer Center, Sci. Center Biol. Res., (Pushchino), 1988.

1763. SHALLOWAY, D., AND DEERING, N. R. Recombinant DNA data management at the restriction and function site level. *Nucl. Acids Res.* **12** (1984), 739-750.
1764. SHALLOWAY, D., SNEDDON, S. F., AND LITTLE, E. K. Microcomputer-based three-dimensional and stereoscopic macromolecular graphics display. *CABIOS* **4** (1988), 193-196.
1765. SHAPIRO, B. A. An algorithm for comparing multiple RNA secondary structures. *CABIOS* **4** (1988), 387-393.
1766. SHAPIRO, B. A., AND LIPKIN, L. E. Nucleic acid morphology: Analysis and synthesis, (in) *Computing in Biological Science*, Geisow, M.J., Barrett, A.N., (eds.), (Elsevier/North-Holland, Amsterdam), 1983.
1767. SHAPIRO, B. A., LIPKIN, L. E., AND MAIZEL, J. An interactive technique for the display of nucleic acid secondary structure. *Nucl. Acids Res.* **10** (1982), 7041-7052.
1768. SHAPIRO, B. A., MAIZEL, J. V., LIPKIN, L. E., CURREY, K., AND WHITNEY, C. Generating non-overlapping displays of nucleic acid secondary structure. *Nucl. Acids Res.* **12** (1984), 75-88.
1769. SHAPIRO, B. A., NUSSINOV, R., LIPKIN, L. E., AND MAIZEL, J. V. Jr. A sequence analysis system encompassing rules for DNA helical distortion. *Nucl. Acids Res.* **14** (1986), 75-86.
1770. SHAPIRO, B. A., NUSSINOV, R., LIPKIN, L. E., AND MAIZEL, J. V. Jr. An interactive dot matrix system for locating potentially significant features in nucleic acid molecules. *J. Biomol. Struct. Dyn.* **4** (1987), 697-706.
1771. SHAPIRO, B. A., AND ZHANG, K. Comparing multiple RNA secondary structures using tree comparison. *CABIOS* **12(4)** (1990), 309-318.
1772. SHAPIRO, M. B. An algorithm for reconstructing protein and RNA sequences. *J. Assoc. Comput. Mach.* **14** (1967), 720-731.
1773. SHAPIRO, M. B., AND SENAPATHY, P. Automated preparation of DNA sequences for publication. *Nucl. Acids Res.* **14** (1986), 65-73.
1774. SHAPIRO, M. B., AND SENAPATHY, P. RNA splice junctions of different classes of eukaryotes: sequence statistics and functional implications in gene expression. *Nucl. Acids Res.* **15** (1987), 7155-7174.
1775. SHARP, P. M. Codon usage and the silent molecular clock: variation among genes and among organisms, (in) *Modelling and Computer Methods in Molecular Biology and Genetics*, Kolchanov, N.A., Ratner, V.A., (eds.), (Abstracts of the International Conference, Novosibirsk), pp.166-167, 1990.
1776. SHARP, P. M., COWE, E., HIGGINS, D. G., SHIELDS, D. C., WOLFE, K. H., AND WRIGHT, F. Codon usage patterns in *Escherichia coli*, *Bacillus subtilis*, *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe*, *Drosophila melanogaster* and *Homo sapiens*; a review of the considerable within-species diversity. *Nucl. Acids Res.* **16** (1988), 8207-8211.
1777. SHARP, P. M., AND DEVINE, K. M. Codon usage and gene expression level in dictyostelium discoideum: highly expressed genes do 'prefer' optimal codons. *Nucl. Acids Res.* **17** (1989), 5029-5039.
1778. SHARP, P. M., AND LI, W. H. Codon usage in regulatory genes in *Escherichia coli* does not reflect selection for 'rare' codons. *Nucl. Acids Res.* **14** (1986), 7737-7749.
1779. SHARP, P. M., AND LI, W. H. An evolutionary perspective on synonymous codon usage in unicellular organisms. *J. Mol. Evol.* **24** (1986), 28-38.
1780. SHARP, P. M., AND LI, W. H. The rate of synonymous substitution in enterobacterial genes is inversely related to codon usage bias. *Mol. Biol. Evol.* **4** (1986), 222-230.
1781. SHARP, P. M., AND LI, W. H. The codon adaptation index-a measure of directional synonymous codon usage bias, and its potential applications. *Nucl. Acids Res.* **15** (1987), 1281-1295.
1782. SHARP, P. M., ROGERS, M. S., AND McCONNELL, D. J. Selection pressures on codon usage in the complete genome of bacteriophage T7. *J. Mol. Evol.* **21** (1985), 150-160.
1783. SHARP, P. M., TUOHY, T. M. F., AND MOSURSKI, K. R. Codon usage in yeast: cluster analysis clearly differentiates highly and lowly expressed genes. *Nucl. Acids Res.* **14** (1986), 5125-5143.
1784. SHAPIRO, M. B., MERRIL, C. R., BRADLEY, D. F., AND MOSIMANN, J. E. Reconstruction of protein and nucleic acid sequences: Alanine transfer ribonucleic acid. *Science*. **150** (1965), 918-921.
1785. SHEN, C. C., AND SHEN, C. K. J. Specificity and flexibility of the recognition of DNA helical structure by eukaryotic topoisomerase I. *J. Mol. Biol.* **212** (1990), 67-78.
1786. SHEPELEV, V. A., BORODOVSKII, M. Yu., AND ALEKSANDROV, A. A. Recognition and classification of protein-coding regions in natural DNA sequences by context features, (in) *Mathematical and Computational Methods in Biology*, (Pushchino), Biomolecular System:83-84, 1987.
1787. SHEPHERD, J. C. W. Method to determine the reading frame of a protein from the purine/pyrimidine genome sequence and its possible evolutionary justification. *Proc. Natl. Acad. Sci. USA* **78** (1981), 1596-1600.
1788. SHEPHERD, J. C. W. Periodic correlations in DNA sequences and evidence suggesting their evolutionary origin in a comma-less genetic code. *J. Mol. Evol.* **17** (1981), 94-102.
1789. SHEPHERD, J. C. W. Fossil remnants of a primeval genetic code in all forms of life? *Trends Biochem. Sci.* **9** (1984), 8-10.
1790. SHEPHERD, J. C. W. Ancient patterns in nucleic acid sequences, (in) *Methods in Enzymology-Molecular Evolution: Computer Analysis of Protein and Nucleic Acid Sequences*, Vol.183, Doolittle, R.F., (ed.), (Academic Press, NY), pp.180-192, 1990.

1791. SHERIDAN, R. P., DIXON, J. S., AND VENKATARAGHAVAN, R. Generating plausible protein folds by secondary structure similarity. *Int. J. Pept. Protein Res.* **25** (1985), 132-143.
1792. SHERMAN, S. A., DEGTEREVA, L. N., RAKOVA, A. A., AND TIMOSHCHUK, T. A. Information and software to operate on the three-dimensional protein structure data base, (in) *Abstracts of: Application of Data Bases and Microcomputers in Molecular Biology*, (Jena), p.54, 1988.
1793. SHESTOPALOV, B. V. Amino acid template useful for alpha-helix-turn- alpha-helix prediction. *FEBS Lett.* **233** (1988), 105-108.
1794. SHESTOPALOV, B. V. Prediction of DNA-recognizing protein supersecondary structure alpha-helix-turn-alpha-helix by a method based on modified Ohlendorf- Anderson-Matthews necessary stereochemical existence requirements. *Mol. Biol. (Mosk.)* **22** (1988), 323-330.
1795. SHIELDS, D. C. Switches in species-specific codon preferences: the influence of mutation biases. *J. Mol. Evol.* **31** (1990), 71-80.
1796. SHIELDS, D. C., AND SHARP, P. M. Synonymous codon usage in *Bacillus subtilis* reflects both translational selection and mutational biases. *Nucl. Acids Res.* **15** (1987), 8023-8040.
1797. SHIELDS, D. C., AND SHARP, P. M. Evidence that mutation rates vary among *drosophila* transposable elements. *J. Mol. Biol.* **207** (1989), 843-846.
1798. SHIELDS, D. C., SHARP, P. M., HIGGINS, D. H., AND WRIGHT, F. "Silent" sites in *Drosophila* genes are not neutral: Evidence of selection among synonymous codons. *Mol. Biol. Evol.* **5** (1988), 704-716.
1799. SHORTRIDGE, R. D., PIRTE, I. L., AND PIRTE, R. M. IBM microcomputer programs that analyze DNA sequences for tRNA genes. *CABIOS* **2** (1986), 13-17.
1800. SHPAER, E. G. The secondary structure of mRNAs from *E. coli*: its possible role in increasing the accuracy of translation. *Nucl. Acids Res.* **13** (1985), 275-288.
1801. SHPAER, E. G. Constraint on codon context in *Escherichia coli* genes. Their possible role in modulating the efficiency of translation. *J. Mol. Biol.* **188** (1986), 555-564.
1802. SHULMAN, M. J., STEINBERG, C. M., AND WESTMORELAND, N. The coding function of nucleotide sequences can be discerned by statistical analysis. *J. Theor. Biol.* **88** (1981), 409-420.
1803. SIBANDA, B. L., AND THORNTON, J. M. Conformation of β hairpins in protein structures: classification and diversity in homologous structures, (in) *Methods in Enzymology-Molecular design and Modeling: Concepts and Applications*, Part A, Vol.202, Langone, J.J., (ed.), (Academic Press, NY), pp.59-82, 1991.
1804. SIBBALD, P. R. Patterns of base usage, nearest neighbour analysis and identification of genes in two completely sequenced chloroplast genomes. *Current Genetics* **13** (1988), 523-530.
1805. SIBBALD, P. R., BANERJEE, S., AND MAZE, J. Constraints on the organization and information properties of DNA sequences. Ph.D. Thesis, University of British Columbia, 1988.
1806. SIBBALD, P. R., BANERJEE, S., AND MAZE, J. Calculating higher order DNA sequence information measures. *J. Theor. Biol.* **136** (1989), 475-483.
1807. SIBBALD, P. R., SOMMERFELDT, H., AND ARGOS, P. Identification of proteins in sequence databases from amino acid composition data. *Analytical Biochemistry* **198** (1991), 330-333.
1808. SIDMAN, K. E., GEORGE, D. G., BARKER, W. C., AND HUNT, L. T. The protein identification resource. *Nucl. Acids Res.* **16** (1988), 1869-1871.
1809. SIKARIS, K., MINASIAN, E., LEACH, S. J., AND FLEGG, R. Computer program designed to predict and plot the secondary structure of proteins. *CABIOS* **5(4)** (1989), 323.
1810. SILVERMAN, B. D., AND LINSKER, R. A measure of DNA periodicity. *J. Theor. Biol.* **118** (1986), 295-300.
1811. SIMOES-PEREIRA, J. M. S., AND ZAMFIRESCU, C. M. Submatrices of non-tree-realizable distance matrices. *Linear Algebra Appl.* **44** (1982), 1-17.
1812. SINGH, A. K., AND OVERBEEK, R. Derivation of efficient parallel programs: An example from genetic sequence analysis. *Int. J. Parallel Programming* **18** (1989), 447-484.
1813. SINGH, A. K., AND OVERBEEK, R. Derivation of efficient parallel programs: An example from genetic sequence analysis, ANL Mathematics and Computer Science Division, preprint MCS-P104-0989, 1989.
1814. SINGH, J., SALDANHA, J., AND THORNTON, J. M. A novel method for the modelling of peptide ligands to their receptors. *Protein Engineering* **4(3)** (1991), 251-261.
1815. SINGHAL, R. P., RAY, R. C., AND DOBBS, L. Computer program for storage and retrieval of the nucleic acid structures. *Comp. Prog. Biomed.* **14** (1982), 277-282.
1816. SIROTKIN, K., AND LOEHR, J. J. Simulation and analysis of physical mapping, 1989. preprint.
1817. SITTIG, D. F., FOULSER, D., CARRIERO, N., MCCORKLE, G., AND MILLER, P. L. A parallel computing approach to genetic sequence comparison: The master-worker paradigm with interworker communication. *Computers Biomed. Res.* **24** (1991), 152-169.
1818. SJORBERG, S., CARLSSON, P., ENERBACK, S., AND BJURSELL, G. A compact, flexible and cheap system for acquiring sequence data from autoradiograms with a digitizer and transferring it to an arbitrary host computer. *CABIOS* **5** (1989), 41-46.
1819. SKALKA, A., BURGI, E., AND HERSEY, A. D. Segmental distribution of nucleotides in the DNA of bacteriophage lambda. *J. Mol. Biol.* **34** (1968), 1-16.
1820. SKLENAR, H. Sequence dependent curvatures of DNA, (in) *Abstracts of: Application of Data Bases and Microcomputers in Molecular Biology*, (Jena), p.58, 1988.

1821. SKLENAR, H., ETCHEBEST, C., AND LAVERY, R. Describing protein structure: a general algorithm yielding complete helicoidal parameters and a unique overall axis. *Proteins* 6 (1989), 46-60.
1822. SKOLINICK, J., AND KOLINSKI, A. Dynamic monte carlo simulations of a new lattice model of globular protein folding, structure and dynamics. *J. Mol. Biol.* 221 (1991), 499-531.
1823. SMETANIC, Ya. S., AND POLOZOV, R. V. On the algorithms for determining the primary structure of biopolymers. *Bull. Math. Biol.* 41 (1979), 1-20.
1824. SMILLIE, F., AND BAINS, W. Repetition structure of mammalian nuclear DNA. *J. Theor. Biol.* 142 (1990), 463-472.
1825. SMITH, D. H., BRUTLAG, D. L., FRIEDLAND, P., AND KEDES, L. H. BIONET: a national computer resource for molecular biology. *Nucl. Acids Res.* 14 (1986), 17-20.
1826. SMITH, G. R. DNA supercoiling: another level for regulating gene expression. *Cell* 24 (1981), 599.
1827. SMITH, H. O., ANNAU, T. M., AND CHANDRASEGARAN, S. Finding sequence motifs in groups of functionally related proteins. *Proc. Natl. Acad. Sci. USA* 87 (1990), 826-830.
1828. SMITH, L. M., SANDERS, J. Z., KAISER, R. J., HUGHES, P., DODD, C., CONNELL, C., HEINER, C., KENT, S. B. H., AND HOOD, L. E. Fluorescence detection and automating DNA sequence analysis. *Nature* 321 (1986), 674-679.
1829. SMITH, M. W. Structure of Vertebrate Genes: A statistical analysis implicating selection. *J. Mol. Biology* 27 (1988), 45-55.
1830. SMITH, R. A finite state machine algorithm for finding restriction sites and other pattern matching applications. *CABIOS* 4(5) (1988), 459-465.
1831. SMITH, R. F., AND SMITH, T. F. Automatic generation of primary sequence patterns from sets of related protein sequences. *Proc. Natl. Acad. Sci.* 87 (1990), 118-122.
1832. SMITH, R. J. The analysis of nucleic acid sequences, (in) **Microcomputers in Biology: A Practical Approach**, Ireland, C.R., Long, S.P., (eds.), (IRL Press, Oxford), pp.151-177, 1984.
1833. SMITH, T. F. Semantic and syntactic patterns in the genetic language, (in) **Biomolecular Data: A Resource in Transition**, Swartz D.G., MacDonell M.T., Colwell R.R., (eds.), (Oxford University Press, NY), pp.211-226, 1989.
1834. SMITH, T. F., AND BURKS, C. Searching for sequence similarities. *Nature* 301 (1983), 194.
1835. SMITH, T. F., GRUSKIN, K., TOLMAN, S., AND FAULKNER, D. V. The molecular biology computer research resource. *Nucl. Acids Res.* 14 (1986), 25-29.
1836. SMITH, T. F., RALPH, W. W., GOODMAN, M., AND CZELUSNIAK, J. Codon usage in vertebrate hemoglobins and its implications. *Mol. Biol. Evol.* 2 (1985), 390-398.
1837. SMITH, T. F., AND WATERMAN, M. S. Comparison of biosequences. *Adv. Appl. Math.* 2 (1981), 482-489.
1838. SMITH, T. F., AND WATERMAN, M. S. Identification of common molecular subsequences. *J. Mol. Biol.* 147 (1981), 195-197.
1839. SMITH, T. F., WATERMAN, M. S., AND BURKS, C. The statistical distribution of nucleic acid similarities. *Nucl. Acids Res.* 13(2) (1985), 645-656.
1840. SMITH, T. F., WATERMAN, M. S., AND FITCH, W. M. Comparative biosequence metrics. *J. Mol. Evol.* 18 (1981), 38-46.
1841. SMITH, T. F., WATERMAN, M. S., AND SADLER, J. R. Statistical characterization of nucleic acid sequence functional domains. *Nucl. Acids Res.* 11 (1983), 2205-2220.
1842. SMITHIES, O., ENGELS, W. R., DEVEREUX, J., SLIGHTOM, J. L., AND SHEN, S. Base substitutions, length differences and DNA strand asymmetries in the human g-gamma and a-gamma fetal globin gene region. *Cell* 26 (1981), 345-353.
1843. SMITHIES, O., GIBSON, D., FANNING, E. M., GOODFLIESH, R. M., GILMAN, J. G., AND BALLANTYNE, D. L. Quantitative procedures for use with the Edman-Begg sequenator. Partial sequences of two unusual immunoglobulin light chains, Rzf and Sac. *Biochemistry* 10 (1971), 4912-4921.
1844. SOBEL, E., AND MARTINEZ, H. M. A multiple sequence alignment program. *Nucl. Acids Res.* 14 (1986), 363-374.
1845. SOBIESKI, J. M., CHEN, K. N., FILIATREAU, J. C., PICKETT, M. H., AND FOX, G. E. 16S rRNA oligonucleotide catalog data base. *Nucl. Acids Res.* 12 (1984), 141-148.
1846. SODERLUND, C. A., SHANMUGAM, P., AND FIELDS, C. A. Automated sequence analysis using the gm system, (in) **Genome Sequencing Conference II**, (abstracts), Hyatt Regency, Hilton Head Island, South Carolina, 1990.
1847. SOLL, D., AND , (eds), R. J. R. **The Applications of Computers to Research on Nucleic Acids II**, Part 1, (IRL Press, Washington, D.C.), 1984.
1848. SOLOVYEV, V. V., KOROLEV, S. V., AND LIM, H. A. A new approach for classification of functional regions of DNA sequences based on fractal representation, FSU-SCRI-91-40, March 1991.
1849. SOLOVYOV, V. V. Informational structure of the functional sites database, (in) **Theoretical Research and Data Banks in Molecular Biology and Genetics** (Abstracts of the 3rd National Conference, Novosibirsk), pp.14-15, 1988.
1850. SOLOVYOV, V. V., ROGOZIN, I. B., SALIKHOVA, A. K., SELEDTSOV, I. A., AND SALAMOV, A. A. "CONTEXT" program package for analysis of structure and function of genetic macromolecules, (in) **Modelling and Computer Methods in Molecular Biology and Genetics**, Kolchanov, N.A., Ratner, V.A., (eds.), (Abstracts of the International Conference, Novosibirsk), pp.219-221, 1990.

1851. SOLOVYOV, V. V., AND SALIKHOVA, A. K. Data base of DNA, RNA and protein functional regions, (in) **Modelling and Computer Methods in Molecular Biology and Genetics**, Kolchanov, N.A., Ratner, V.A., (eds.), (Abstracts of the International Conference, Novosibirsk), pp.251-252, 1990.
1852. SOLOVYOV, V. V., SALIKHOVA, A. K., AND ROGOZIN, I. B. A system for genetic texts functional diagnostics, (in) **Computer Analysis of Structure, Function, and Evolution of Genetic Macromolecules, Intellectualization Problems**, Kolchanov, N.A., (ed.), (Novosibirsk), pp.12-40, 1989.
1853. SOLOVYOV, V. V., SELEDTSOV, I. A., AND SALIKHOVA, A. K. System for revealing of complicated structural-functional organization of genetic texts, (in) **Modelling and Computer Methods in Molecular Biology and Genetics**, Kolchanov, N.A., Ratner, V.A., (eds.), (Abstracts of the International Conference, Novosibirsk), pp.55-56, 1990.
1854. SOO, V. A Qualitative Matching Scheme For Postulating Enzyme Kinetic Models and Experimental Conditions: Reasoning With Constraints, Ph.D. Thesis, Rutgers University, 1987.
1855. SOURDIS, J., AND KRIMBAS, C. Accuracy of phylogenetic trees estimated from DNA sequence data. *Mol. Biol. Evol.* 4 (1987), 159-166.
1856. SOURDIS, J., AND NEI, M. Relative efficiencies of the maximum parsimony and distance-matrix methods in obtaining the correct phylogenetic tree. *Mol. Biol. Evol.* 5 (1988), 298-311.
1857. SOUTHERN, E. M. Measurement of DNA length by gel electrophoresis. *Anal. Biochem.* 100 (1979), 319-323.
1858. SOUTHERN, E. M. Calculation of size from mobility, (in) **Methods in Enzymology**, Vol.68, (Academic Press, NY), pp.161-162, 1980.
1859. SPITZNER, J. R., CHUNG, I. K., AND MULLER, M. T. Eukaryotic topoisomerase II preferentially cleaves alternating purine-pyrimidine repeats. *Nucl. Acids Res.* 18 (1990), 1-11.
1860. SPITZNER, J. R., AND MULLER, M. T. Application of a degenerate consensus sequence to quantify recognition sites by vertebrate DNA topoisomerase II. *J. Mol. Recognition* 2 (1989), 63-74.
1861. SPOUGE, J. L. Improving sequence-matching algorithms by working from both ends. *J. Mol. Biol.* 181 (1985), 137-138.
1862. SPOUGE, J. L. Fast optimal alignment I. Single indel wieghts, *CABIOS*, submitted, 1989.
1863. SPOUGE, J. L. Speeding up dynamic programming algorithms for finding optimal lattice paths . *SIAM J. Appl. Math* 49(5) (1989), 1552-1566.
1864. SPOUGE, J. L. Fast optimal alignment. *CABIOS* 7(1) (1991), 1-7.
1865. SPRENGART, M. L., FATSCHER, H. P., AND FUCHS, E. The initiation of translation in *E. coli*: apparent base pairing between the 16s rRNA and downstream sequences of the mRNA. *Nucl. Acids Res.* 18 (1990), 1719-1723.
1866. SPRINZL, M., HARTMANN, T., MEISSNER, F., MOLL, J., AND VORDERWULBECKE, T. Compilation of tRNA sequences. *Nucl. Acids Res.* 15 (1987), r53-r188.
1867. SPRIZHITSKY, Y. A. The new format for data base of the biopolymer primary structures, (in) **Abstracts of: Application of Data Bases and Microcomputers in Molecular Biology**, (Jena), p.17, 1988.
1868. SPRIZHITSKY, Y. A., NECHIPURENKO, Y. D., ALEXANDROV, A. A., AND VOLKENSTEIN, M. Statistical analysis of nucleotide runs in coding and noncoding DNA sequences. *J. Biomol. Struct. Dyn.* 6 (1988), 345-358.
1869. SPRIZHITSKY, Yu. A., GELFAND, M. S., AND KALAMBET, Yu. A. A data base of functional signal prediction models, (in) **Human Genome-90**, Bayev, A.E. (ed.), (Abstracts of the 1st National Conference, Pereslavl-Zalesky), (Moscow), pp.210-211, 1990.
1870. SPRIZHITSKY, Yu. A., AND KOPYLOV, V. The SOS system of *Escherichia coli* in the regulation of bacteriophage lambda development. *FEBS Lett.* 160 (1983), 7-10.
1871. SPRIZHITSKY, Yu. A., NECHIPURENKO, Yu. A., ALEXANDROV, A. A., AND VOLKENSTEIN, M. V. Statistical analysis of nucleotide runs in coding and noncoding DNA sequences. *Mol. Biol.* 22 (1990), 338-356.
1872. SREDNEVA, V. V., ROGOZIN, I. B., AND KOLCHANOV, N. A. Method of statistical weights for analysis of functional characteristics of nucleotide sequences, (in) **Theoretical Research and Data Banks in Molecular Biology and Genetics**, (Abstracts of the 3rd National Conference, Novosibirsk), pp.35-36, 1988.
1873. STADEN, R. . Methods for discovering novel motifs in nucleic acid sequences. *CABIOS* 5(1) (1989), 293-298.
1874. STADEN, R. Sequence data handling by computer. *Nucl. Acids Res* 4(11) (1977), 4037-4051.
1875. STADEN, R. Further procedures for sequence analysis by computer. *Nucl. Acids Res.* 5(3) (1978), 1013-1015.
1876. STADEN, R. A strategy of DNA sequencing in employing computer programs. *Nucl. Acids. Res.* 6 (1979), 2601-2610.
1877. STADEN, R. A computer program to search for tRNA genes. *Nucl. Acids Res.* 8 (1980), 817-825.
1878. STADEN, R. A new computer method for the storage and manipulation of DNA gel reading data. *Nucl. Acids Res.* 8(16) (1980), 3673-3694.
1879. STADEN, R. Automation of the computer handling of gel reading data produced by the shotgun method of DNA sequencing. *Nucl. Acids Res.* 10(15) (1982), 4731-4751.

1880. STADEN, R. An interactive graphics program for comparing and aligning nucleic acid and amino acid sequences. *Nucl. Acids Res.* **10**(9) (1982), 2951-2961.
1881. STADEN, R. Computer methods to aid the determination and analysis of DNA sequences. *Biochem. Soc. Trans.* **12** (1984), 1005-1008.
1882. STADEN, R. A computer program to enter DNA gel reading data into a computer. *Nucl. Acids Res.* **12** (1984), 499-503.
1883. STADEN, R. Graphic methods to determine the function of nucleic acid sequences. *Nucl. Acids Res.* **12** (1984), 521-538.
1884. STADEN, R. Measurements of the effect that coding for a protein has on a DNA sequence and their use for finding genes. *Nucl. Acids Res.* **12** (1984), 551-567.
1885. STADEN, R. Computer methods to locate genes and signals in nucleic acid sequences, (in) **Genetic Engineering, Principle and Methods**, Setlow J.K., Hollaender A., (eds.), 7:67-114, (Plenum Press, NY), 1985.
1886. STADEN, R. Computer handling of DNA sequencing projects, (in) **Nucleic Acid and Protein Sequence Analysis: A Practical Approach**, Bishop, M.J., Rawlings, C.J., (eds.), (IRL Press, Oxford), pp.359-385, 1987.
1887. STADEN, R. Methods to define and locate patterns of motifs in sequences. *CABIOS* **4**(1) (1988), 53-60.
1888. STADEN, R. Methods for calculating the probabilities of finding patterns in sequences. *CABIOS* **5**(2) (1989), 89-96.
1889. STADEN, R. Finding protein coding regions in genomic sequences, (in) **Methods in Enzymology-Molecular Evolution: Computer Analysis of Protein and Nucleic Acid Sequences**, Vol.183, Doolittle, R.F., (ed.), (Academic Press, NY), pp.163-180, 1990.
1890. STADEN, R. Searching for patterns in protein and nucleic acid sequences, (in) **Methods in Enzymology-Molecular Evolution: Computer Analysis of Protein and Nucleic Acid Sequences**, Vol.183, Doolittle, R.F., (ed.), (Academic Press, NY), pp.193-211, 1990.
1891. STADEN, R., AND McLACHLAN, A. D. Codon preference and its use in identifying protein coding regions in long DNA sequences. *Nucl. Acids Res.* **10** (1982), 151-156.
1892. STANFORD, D. R., PENNISTON, J. T., AND WIEBEN, E. D. Database pollution. *Nature* **342** (1989), 24-24.
1893. STATES, D. J. Evaluation the reliability for direct analysis of imperfect sequences, (in) **Genome Sequencing Conference II**, (abstracts), Hyatt Regency, Hilton Head Island, South Carolina, 1990.
1894. STATES, D. J., AND BOTSTEIN, D. Molecular sequence accuracy and the analysis of protein coding regions. *Proc. Natl. Acad. Sci. USA* **88** (1991), 5518-5522.
1895. STATES, D. J., AND BOTSTEIN, D. Molecular sequence data accuracy and the analysis of protein coding regions. *Proc. Natl. Acad. Sci. USA* **88** (1991), 5518-5522.
1896. STEFIK, M. Inferring DNA structures from segmentation data. *Artif. Intell.* **11** (1978), 85-114.
1897. STEFIK, M. J. Planning with Constraints, Ph.D. Thesis, Stanford University Computer Science Department, 1980. See also Stanford University Computer Science Department reports HPP-80-2, STAN-CS-80-784.
1898. STERNBERG, M. J. E. The analysis and prediction of protein structure, (in) **Computing in Biological Science**, Geisow, M.J., Barrett, A.N., (eds.), (Elsevier/North-Holland, Amsterdam), 1983.
1899. STERNBERG, M. J. E., AND COHEN, F. E. Prediction of the secondary and tertiary structures of interferon from four homologous amino acid sequences. *Int. J. Biol. Macromol.* **4** (1982), 137-144.
1900. STERNBERG, M. J. E., AND ISLAM, S. A. A relational database of protein structure. *Biochem. Soc. Trans.* **17** (1989), 845-847.
1901. STERNBERG, M. J. E., AND THORNTON, J. M. Prediction of protein structure from amino acid sequence. *Nature* **271** (1978), 15-20.
1902. STERNBERG, M. J. E., AND THORNTON, J. M. Prediction of protein structure from amino acid sequence. *Biochem. Soc. Trans.* **6** (1978), 1119-1123.
1903. STOCKWELL, P. A. A large database DNA sequence handling program with generalized searching specifications. *Nucl. Acids Res.* **10** (1982), 115-125.
1904. STOEHR, P. J., AND OMOND, R. A. The EMBL network file server. *Nucl. Acids Res.* **17** (1989), 6763-6764.
1905. STOLORZ, P., LAPEDES, A., AND XIA, Y. Predicting protein secondary structure using neural net and statistical methods. Tech. Rep. LA-UR-91-15, Los Alamos National Laboratory, Theoretical Division, MS B213, Los Alamos National Laboratory, Los Alamos, NM, 87545, 1991.
1906. STONE, B. N., GRIESINGER, G. L., AND MODELEVSKY, J. L. PLASMAP: An interactive computational tool for storage, retrieval, and device-independent graphic display of conventional restriction maps. *Nucl. Acids Res.* **12** (1984), 465-471.
1907. STONE, T. W., AND POTTER, K. N. A DNA analysis program designed for computer novices working in an industrial research environment. *Nucl. Acids Res.* **12** (1984), 367-378.
1908. STONE, T. W., AND POTTER, K. N. Methylation blockage and other improvements to a comprehensive DNA analysis program. *Nucl. Acids Res.* **14** (1986), 255-264.
1909. STORMO, G. Identifying regulatory sites from DNA sequence data, (in) **Structure and Methods**, Proc. of the 6th Conversation, State University of New York, Albany, Sarma, R.H., Sarma, M.H., (eds.) Vol.1:Human Genome Initiative and DNA Recombination, 1989.

1910. STORMO, G. D. Consensus patterns in DNA, preprint.
1911. STORMO, G. D. Identifying coding sequences, (in) **Nucleic Acid and Protein Sequence Analysis: A Practical Approach**. Bishop, M.J., Rawlings, C.J., (eds.), (IRL Press, Oxford), pp.359-385, 1987.
1912. STORMO, G. D. Computer methods for analyzing sequence recognition of nucleic acids. *Annu. Rev. Biophys. Biophys. Chem.* 17 (1988), 241-263.
1913. STORMO, G. D. Consensus patterns in DNA, (in) **Methods in Enzymology-Molecular Evolution: Computer Analysis of Protein and Nucleic Acid Sequences**, Vol.183, Doolittle, R.F., (ed.), (Academic Press, NY), pp.211-221, 1990.
1914. STORMO, G. D., AND HARTZELL, G. W. III. Identifying protein-binding sites from unaligned DNA fragments. *Proc. Natl. Acad. Sci.* 86 (1989), 1183-1187.
1915. STORMO, G. D., SCHNEIDER, T. D., AND GOLD, L. Quantitative analysis of the relationship between nucleotide sequence and functional activity. *Nucl. Acids Res.* 14(16) (1986), 6661-6679.
1916. STORMO, G. D., SCHNEIDER, T. D., GOLD, L., AND EHRENFEUCHT, A. Use of the Perceptron algorithm to distinguish translational initiation sites in *E. coli*. *Nucl. Acids Res.* 10(9) (1982), 2997-3011.
1917. STORMO, G. D., SCHNEIDER, T. D., AND GOLD, L. M. Characterization of translational initiation sites in *E. coli*. *Nucl. Acids Res.* 10 (1982), 2971-2996.
1918. STUBER, K. Visualisation of nucleic acid sequence structural information. *CABIOS* 1 (1985), 35-42.
1919. STUBER, K. Nucleic acid secondary structure prediction and display. *Nucl. Acids Res.* 14 (1986), 317-326.
1920. STUBER, K., AND SPANIER, K. SEQIN-ST and CODFREG: a full screen sequence editor and a codon usage analysis program for the Atari-ST. *Nucl. Acids Res.* 16 (1988), 1821-1827.
1921. STUDIER, J. A., AND KEPPLER, K. J. A note on the neighbor-joining algorithm of Saitou and Nei. *Mol. Biol. Evol.* 5 (1988), 729-731.
1922. STUDNICKA, G., RAHN, I., CUMMINGS, W., AND SALSER, W. A. Computer methods for predicting the secondary structure of single stranded RNA. *Nucl. Acids Res.* 5 (1978), 3365-3387.
1923. STUDNICKA, G. M. Quantitative computer analysis of signal sequence homologies in DNA. *CABIOS* 2 (1986), 269-275.
1924. STUDNICKA, G. M. Nucleotide sequence homologies in control regions of prokaryotic genomes. *Gene* 58 (1987), 45-57.
1925. STUDNICKA, G. M. *Escherichia coli* promoter -10 and -35 homologies correlate with binding and isomerization kinetics. *Biochemical Journal* 252 (1988), 825-831.
1926. STULICH, R. Computer generation of a protein sequence data base, (in) **Abstracts of: Application of Data Bases and Microcomputers in Molecular Biology**, (Jena), p.43, 1988.
1927. STULICH, R., AND ROHDE, K. EMBOPRO-an automatically generated protein sequence database. *CABIOS* 5(1) (1989), 15-18.
1928. Subcommittee on Human Genome of the HERAC for the U.S. Department of Energy, Office of Energy Research, Office of Health and Environmental Research. Report on the Human Genome Initiative for the Office of Health and Environmental Research, April, 1987.
1929. SUBBA, R. J., GEEVAN, C. P., AND SUBBA, R. G. Significance of the information content of DNA in mutations and evolution. *J. Theor. Biol.* 96 (1982), 571-577.
1930. SUBBIAH, S., AND HARRISON, S. C. A method for multiple sequence alignment with gaps. *J. Mol. Biol.* 209 (1989), 539-548.
1931. SUBOCH, G. M., AND SPRIZHITSKY, Y. A. Statistical significance of some complex structures in DNA, (in) **Abstracts of: Application of Data Bases and Microcomputers in Molecular Biology**, p.61, 1988.
1932. SUBOCH, G. M., AND SPRIZHITSKY, Y. A. Splice junctions of human genes: statistical analysis and discrimination, (in) **Modelling and Computer Methods in Molecular Biology and Genetics**, Kolchanov, N.A., Ratner, V.A., (eds.), (Abstracts of the International Conference, Novosibirsk), p.32, 1990.
1933. SUBOCH, G. M., AND SPRIZHITSKY, Y. A. Statistical significance of some complex nucleotide combinations: a comparison of DNA models. *CABIOS* 6 (1990), 43-48.
1934. SUBOCH, G. M., SPRIZHITSKY, Yu. A., AND ALEXANDROV, A. A. Occurrence of homopurine-homopyrimidine mirror repeats in natural DNA. *Studia Biophysica* 129 (1989), 201-208.
1935. SULSTON, J., MALLETT, F., DURBIN, R., AND HORSNELL, T. Image analysis of restriction enzyme fingerprint autoradiograms. *CABIOS* 5(2) (1989), 101-106.
1936. SULSTON, J., MALLETT, F., STADEN, R., DURBIN, R., HORSNELL, T., AND COULSON, A. Software for genome mapping by fingerprinting techniques. *CABIOS* 4(1) (1988), 125-132.
1937. Summer Program on Statistical Computation, Inst. Math. and Its Appl., Univ. Minn. Minneapolis, MN., (Overheads only).
1938. SUMMERS, N. L., CARLSON, W. D., AND KARPLUS, M. Analysis of side-chain orientations in homologous proteins. *J. Mol. Biol.* 196 (1987), 175-198.
1939. SUMMERS, N. L., AND KARPLUS, M. Modeling of side chains, loops, and insertions in proteins, (in) **Methods in Enzymology-Molecular design and Modeling: Concepts and Applications**, Part A, Vol.202, Langone, J.J., (ed.), (Academic Press, NY), pp.156-204, 1991.
1940. SUMMERS, D. W. Knot theory, statistics, and DNA. *Kem. Ind.* 35(12) (1986), 657-661.

1941. SUMNERS, D. W. The role of knot theory in DNA research, (in) *Geometry and Topology: Manifolds, Varieties, and Knots*, McCrory, C., Shifrin, T., (eds.), (Marcel Dekker, Inc., NY), pp.297-318, 1987.
1942. SUMNERS, D. W. Untangling DNA. *The Mathematical Intelligencer* 12(3) (1990), 71-80.
1943. SUNDAY, D. M. A very fast substring search algorithm. *Commun. ACM.* 3 (1990), 132-142.
1944. SURGUCHOV, A. P. The apolipoprotein gene family organization of upstream elements and regulation of gene expression. *Biomedical Science* 1 (1990), 344-353.
1945. SUTHERLAND, R. A., AND PARTIS, M. D. Using fragment lengths for incomplete digestion by multiply cleaving enzymes to map antibody binding sites on a protein. *CABIOS* 8(1) (1992), 1-4.
1946. SVED, J., AND BIRD, A. The expected equilibrium of the CpG dinucleotide in vertebrate genomes under a mutation model. *Proc. Natl. Acad. Sci. USA* 87 (1990), 4692-4696.
1947. SVED, J., AND COWAN, R. An analysis of the heterogeneity of codon usage between different human genes, (in) *Molecular Evolution*, (UCLA Symposia on Molecular and Cellular Biology) Clegg, M.T., Obrien, S.J., (eds.), (Wiley-Liss, New Series), Vol.122, p.297, 1990.
1948. SWANSON, R. A unifying concept for the amino acid code. *Bull. Math. Biol.* 46 (1984), 187-204.
1949. SWANSON, R. A vector representation for amino acid sequences. *Bull. Math. Biol.* 46 (1984), 623-639.
1950. SWEET, R. M., AND EISENBERG, D. Correlation of sequence hydrophobicities measures similarity in three-dimensional protein structure. *J. Mol. Biol.* 171 (1983), 479-488.
1951. SWINBANKS, D. Genome use for computers. *Nature* 345 (1985), 466-467.
1952. TAJIMA, F. Determination of window size for analyzing DNA sequences. *J. Mol. Evol.* 33 (1991), 470-473.
1953. TAJIMA, F., AND NEI, M. Estimation of evolutionary distance between nucleotide sequences. *Mol. Biol. Evol.* 3 (1984), 269-285.
1954. TAJIMA, K. Multiple DNA and protein sequence alignment on a workstation and a supercomputer. *CABIOS* 4(4) (1988), 467-471.
1955. TAKAHASHI, H., ITOH, N., AMANO, T., AND YAMASHITA, A. A spelling correction method and its application to an OCR system. *Pattern Recognition* 23 (1990), 363-377.
1956. TAKAHASHI, K., VIGNERON, M., MATTHES, H., WILDEMAN, A., ZENKE, M., AND CHAMBON, P. Requirement of stereospecific alignments for initiation from the simian virus 40 early promoter. *Nature* 319 (1986), 121-126.
1957. TAKAHATA, N., AND TAJIMA, F. Sampling errors in phylogeny. *Mol. Biol. Evol.* 8(4) (1991), 494-502.
1958. TAKETOMI, H., KANO, F., AND GO, N. The effect of amino acid substitution on protein-folding and unfolding transition studied by computer simulation. *Biopolymers* 27(4) (1988), 527-559.
1959. TAN, R. K. Z., AND HARVEY, S. C. Molecular mechanics model of supercoiled DNA. *J. Mol. Biol.* 205(3) (1989), 573-591.
1960. TAN, R. K. Z., PRABHAKARAN, M., TUNG, C. -S., AND HARVEY, S. C. AUGUR: a program to predict, display and analyze the tertiary structure of B-DNA. *CABIOS* 4 (1988), 147-151.
1961. TANAKA, S., AND SCHERAGA, H. S. Statistical mechanical treatment of protein conformation. A four state model for specific-sequence copolymers of amino acids. *Macromolecules* 9 (1976), 812-824.
1962. TARNIO, J., AND UKKONEN, E. A greedy approximation algorithm for constructing shortest common superstrings. *Theoretical Computer Science* 57 (1988), 131-145.
1963. TATENO, Y. A method for molecular phylogeny construction by direct use of nucleotide sequence data. *J. Mol. Evol.* 30 (1990), 85-93.
1964. TATENO, Y., NEI, M., AND TAJIMA, F. Accuracy of estimated phylogenetic trees from molecular data. I. Distantly related species. *J. Mol. Evol.* 18 (1982), 387-404.
1965. TAUTZ, D., TRICK, M., AND DOVER, G. A. Cryptic simplicity in DNA is a major source of genetic variation. *Nature* 322 (1986), 652-656.
1966. TAVARE, S., AND GIDDINGS, B. W. Some statistical aspects of the primary structure of nucleotide sequences, (in) *Mathematical Methods for DNA Sequences*, Waterman, M., (ed.), (CRC Press, Boca Raton), 1989.
1967. TAVARE, S., AND SONG, B. Codon preference and primary sequence structure in protein-coding regions. *Bull. Math. Biol.* 51(1) (1989), 95-115.
1968. TAYLOR, P. A fast homology program for aligning biological sequences. *Nucl. Acids Res.* 12(1) (1984), 447-455.
1969. TAYLOR, P. A computer program for translating DNA sequences into protein. *Nucl. Acids Res.* 14 (1986), 437-441.
1970. TAYLOR, P. A program for drawing evolutionary trees. *CABIOS* 4 (1988), 441-443.
1971. TAYLOR, P., ROSENBERG, P., AND SAMSONOVA, M. G. A new method for finding long consensus patterns in nucleic acid sequences. *CABIOS* 7(4) (1991), 495-500.
1972. TAYLOR, W. R. The classification of amino acid conservation. *J. Theor. Biol.* 119 (1986), 205-218.
1973. TAYLOR, W. R. Identification of protein sequence homology by consensus template alignment. *J. Mol. Biol.* 188 (1986), 233-258.
1974. TAYLOR, W. R. Toward a practical grammar of protein structure: protein structure prediction by template fitting, (in) *Computer Graphics and Molecular Modeling*, Fletterick R., Zoller M., (eds.), (Cold Spring Harbor Laboratory, NY), pp. 77-84, 1986.

1975. TAYLOR, W. R. Multiple sequence alignment by a pairwise algorithm. *CABIOS* **3**(2) (1987), 81-87.
1976. TAYLOR, W. R. A flexible method to align large numbers of biological sequences. *J. Mol. Evol.* **28** (1988), 161-169.
1977. TAYLOR, W. R. Pattern matching methods in protein sequence comparison and structure prediction. *Protein Eng.* **2**(2) (1988), 77-86.
1978. TAYLOR, W. R. Hierarchical method to align large numbers of biological sequences, (in) *Methods in Enzymology-Molecular Evolution: Computer Analysis of Protein and Nucleic Acid Sequences*, Vol.183, Doolittle, R.F., (ed.), (Academic Press, NY), pp.456-474, 1990.
1979. TAYLOR, W. R., AND ORENGO, C. A. A holistic approach to protein structure alignment. *Protein Engineering* **2** (1989), 505-519.
1980. TAYLOR, W. R., AND ORENGO, C. A. Protein structure alignment. *J. Mol. Biol.* **208** (1989), 1-22.
1981. TAYLOR, W. R., AND THORNTON, J. M. Prediction of super-secondary structure in proteins. *Nature* **301** (1983), 540-542.
1982. TAYLOR, W. R., AND THORNTON, J. M. Recognition of super-secondary structure in proteins. *J. Mol. Biol.* **173** (1984), 487-514.
1983. TEUFEL, B., AND SCHMIDT, S. Full text retrieval based on syntactic similarities. *Information Systems* **13** (1988), 65-70.
1984. TEWARI, N., AND WAGH, M. D. Bit-sequential array for pattern matching. *Proceedings of the IEEE* **74** (1986), 1465-1466.
1985. THANARAJ, T. A., KOLASKAR, A., AND PANDIT, M. W. An extension of the graph theoretical approach to predict the secondary structure of large RNAs: The complex of 16S and 23S rRNAs from *E. coli*. as a case study. *CABIOS* **5**(3) (1989), 211-218.
1986. THANARAJ, T. A., AND PANDIT, M. W. Translation-initiation promoting site on transcripts of highly expressed genes from *Saccharomyces cerevisiae* and the role of hairpin stems to position the site near the initiation codon. *J. Biomol. Struct. Dyn.* **7** (1990), 1279-1289.
1987. THORNTON, J. M., EDWARDS, M. S., TAYLOR, W. R., AND BARLOW, D. J. Location of "continuous" antigenic determinants in the protruding regions of proteins. *EMBO J.* **5** (1986), 409-413.
1988. THORNTON, J. M., AND GARDNER, S. P. Protein motifs and data-base searching. *Trends Biochem. Sci.* **14** (1989), 300-304.
1989. TICHER, A., AND GRAUR, D. Nucleic acid composition, codon usage and the rate of synonymous substitution in protein-coding genes. *J. Mol. Evol.* **28** (1989), 286-298.
1990. TINOCO, I., BORER, P. N., DENGLER, B., LEVINE, M. D., UHLENBECK, O. C., CROTHERS, D. M., AND GRALLA, J. Improved estimation of secondary structure in ribonucleic acids. *Nature New Biol.* **246** (1973), 40-41.
1991. TOHA, J. C., AND SOTO, M. A. Analogous representation of the genetic code in terms of the virtual base "e" and codons of triplets. *Z. Naturforsch. Sect. C - J. Biosci.* **44** (1989), 1035.
1992. TOLSTOSHEV, C. M., AND BLAKESLEY, R. W. RSITE: A computer program to predict the recognition structure of a restriction enzyme. *Nucl. Acids Res.* **10** (1982), 1-17.
1993. TOLSTOSHEV, C. M., JELTSCH, J. M., FRITZ, R., AND OUDET, P. A DNA recombinant database management system. *Nucl. Acids Res.* **11** (1983), 4611-4627.
1994. TOLSTOSHEV, C. M., MATTHES, H. W. D., AND OUDET, P. Computer management of oligonucleotide synthesis on cellulose filters. *Nucl. Acids Res.* **14** (1986), 405-415.
1995. TOMBOULIAN, S. Introduction to a system for implementing neural net connections on SIMD architectures, (in) *Neural Information Processing Systems*, Dana Z. Anderson, (ed.), (American Institute Physics, NY), 1988.
1996. TOPHAM, C. M., SALIH, E., FRAZAO, C., KOWLESSUR, D., OVERINGTON, J. P., THOMAS, M., BROCKLEHURST, S. M. PATEL, M., THOMAS, E. W., AND BROCKLEHURST, K. Structure-function relationships in the cysteine proteinases actininidin, papain and papaya proteinase Ω. *J. Biochem.* **280** (1991), 79-92.
1997. TORNEY, D. C., BURKS, C., DAVISON, D., AND SIROTKIN, K. M. Computation of d^2 -A measure of sequence dissimilarity, (in) *The Interface Between Computational Science and Nucleic Acid Sequencing*, SFI Studies in the Sciences of Complexity, Bell,G., Marr, T., (eds.), (Addison-Wesley, Reading, Massachusetts), 1989.
1998. TORNEY, D. C., BURKS, C., DAVISON, D., AND SIROTKIN, K. M. Computation of d^2 -A measure of sequence dissimilarity, (in) *Computers and DNA*, SFI Studies in the Sciences of Complexity. Vol.VI, Bell, G., Marr, T., (eds.), (Addison-Wesley, MA), 1989.
1999. TORNEY, D. C., WHITTAKER, C. C., WHITE, S. W., AND SCHENK, K. R. Computational methods for physical mapping of chromosomes, (in) *Proceedings of Conference on Electrophoresis, Supercomputing and the Human Genome*, Tallahassee, FL, April 10-13, 1990.
2000. TRAMONTANO, A., AND MACCHIATO, M. F. Probability of coding of a DNA sequence: an algorithm to predict translated reading frames from their thermodynamic characteristics. *Nucl. Acids Res.* **14** (1986), 127-135.
2001. TRAMONTANO, A., AND MACCHIATO, M. F. A transportable interactive package for the statistical analysis and handling of sequence data. *Comput. Biol. Med.* **18**(2) (1988), 113-122.

2002. TRAUT, T. W. Do exons code for structural or functional units in proteins? *Proc. Natl. Acad. Sci. USA* **85** (1988), 2944–2948.
2003. TRAVERS, A. A. Conserved features of coordinately regulated *E. coli* promoters. *Nucl. Acids Res.* **12** (1984), 2605–2618.
2004. TRAVERS, A. A. DNA bending and nucleosome positioning. *Trends Biochem. Sci.* **12** (1987), 108–112.
2005. TRAVERS, A. A. Protein-induced DNA bending. *Nucl. Acids Mol. Biol.* **2** (1988), 136–148.
2006. TRIFONOV, E. N. Sequence-dependent deformational anisotropy of chromatin DNA. *Nucl. Acids Res.* **8** (1980), 4041.
2007. TRIFONOV, E. N. Topological paradox of minichromosomes, (in) **Structural Aspects of Recognition and Assembly in Biological Macromolecules**, M. Balaban, J. Sussman, W. Traub, A. Yonath, (eds.), (Balaban ISS, Rehovot-Philadelphia), 1981.
2008. TRIFONOV, E. N. Construction of an algorithm for locating splicing junctions. *CODATA Bull.* **56** (1983), 21–26.
2009. TRIFONOV, E. N. Sequence-dependent variations of B-DNA structure and protein-DNA interaction. *Cold Spring Harbor Symp. Quant. Biol.* **47** (1983), 271–278.
2010. TRIFONOV, E. N. Construction of an algorithm for locating splicing junctions. *CODATA Bull.* **56** (1985), 21–26.
2011. TRIFONOV, E. N. Construction of an algorithm for locating splicing junctions, (in) **The Role of Data in Scientific Progress**, Glaeser, P.S., (ed.), (Elsevier, North Holland), pp.119-124, 1985.
2012. TRIFONOV, E. N. Imperfect complementarity and RNA structure, (in) **Proc. Int. Symp. Biomol. Struct. Interactions**, 8, pp.781-790, 1985.
2013. TRIFONOV, E. N. Curved DNA. *CRC Crit. Rev. Biochem.* **19** (1986), 89.
2014. TRIFONOV, E. N. Translation framing code and frame-monitoring mechanism as suggested by the analysis of mRNA and 16S rRNA nucleotide sequences. *J. Mol. Biol.* **194** (1987), 643–652.
2015. TRIFONOV, E. N. Codes of nucleotide sequences. *Math. Bio. Sci.* **90** (1988), 507–517.
2016. TRIFONOV, E. N. Nucleotide sequences as a language: Morphological classes of words, (in) **Classification and Related Methods of Data Analysis**, Bock, H.H., (ed.), (Elsevier Science/North-Holland, Amsterdam), pp.57-64, 1988.
2017. TRIFONOV, E. N. Making sense of the human genome, (in) **Structure and Methods**, (Proc. of the 6th Conversation, State University of New York, Albany), Salama, R.H., Arma, M.H., (eds.), Vol.1: Human Genome Initiative and DNA Recombination, 1989.
2018. TRIFONOV, E. N. Searching for codes in the sequences, (in) **Biomolecular Data: A Resource in Transition**, Swartz, D.G., MacDoneel, M.T., Colwell, R.R., (eds.), (Oxford University Press, NY), pp.199–210, 1989.
2019. TRIFONOV, E. N. Viewpoint: Multiple codes of nucleotide sequences. *Bull. Math. Biol.* **51**(4) (1989), 417–432.
2020. TRIFONOV, E. N., AND BOLSHOI, G. Open and closed 5S rRNA are the only two universal structures encoded in the nucleotide sequences. *J. Mol. Biol.* **169** (1983), 1–13.
2021. TRIFONOV, E. N., AND BRENDL, V. **Gnomic, A Dictionary of Genetic Codes**, Balaban Publishers, Philadelphia, 1986.
2022. TRIFONOV, E. N., KONOPKA, A. N., AND JOVIN, T. M. Unusual frequencies of certain alternating purine-pyrimidine runs in natural DNA sequences: relation to Z-DNA. *FEBS Lett.* **185** (1985), 197–202.
2023. TRIFONOV, E. N., AND LAGUNEZ, J. Translation framing code, (in) **Modelling and Computer Methods in Molecular Biology and Genetics**, Kochanov, N.A., Ratner, V.A., (eds.), (Abstracts of the International Conference, Novosibirsk), p.44, 1990.
2024. TRIFONOV, E. N., AND MENGERITSKY, G. Bent DNA in chromatin versus free curved DNA, (in) **Structure Expression**, Olson, W.K., Salama, M.H., Salama, R.H. and Sundaralingam, M., (eds.), (Adenine Press, NY), 1987.
2025. TRIFONOV, E. N., AND SUSSMAN, J. L. The pitch of chromatin DNA is reflected in its nucleotide sequence. *Proc. Natl. Acad. Sci.* **77** (1980), 3816.
2026. TRIFONOV, E. N., TAN, R. K. Z., AND HARVEY, S. C. Static persistence length of DNA, (in) **Structure Expression**, Vol. 3: DNA Bending and Curvature. Olson, W.K., Salama, M.H., Salama, R.H., Sundaralingam, M., (eds.), (Adenine Press, NY), 1987.
2027. TRIFONOV, E. N., AND ULANOVSKY, L. E. Inherently curved DNA and its structural elements, (in) **Unusual DNA Structures**, Proceedings of the First Gulf Shores Symposium, April 6–8, Wells, R.D., Harvey, S.C., (eds.), (Springer-Verlag, NY), pp.173–187, 1987.
2028. TSUR, S., OLKEN, F., AND NAOR, D. Deductive database for genomic mapping. Microelectronics and Computer Technology Corporation Technical report No.STP-352-90, 1990.
2029. TUERK, C., GAUSS, P., THERMES, C., GROEBE, D. R., GAYLE, M., GUILD, N., STORMO, G., D'AUBENTON CARAFÀ, Y., UHLENBECK, O. C., TINOCO, I. Jr., BROODY, E. N., AND GOLD, L. CUUCGG hairpins: Extraordinary stable RNA secondary structures associated with various biochemical processes. *Proc. Natl. Acad. Sci. USA* **85** (1988), 1364–1368.
2030. TUFFERY, P., DESSEN, P., MUGNIER, C., AND HAZOUT, S. Restriction map construction using complete sentences compatibility algorithm. *CABIOS* **4**(1) (1988), 103–110.

2031. TUNG, C. -S., GARCIA, A. E., MARR, T. M., MISHRA, S. K., AND BURKS, C. Measures of DNA Curvature, (in) *Structure Expression*, W.K. Olson, M.H. Sharma, R.H. Sarma, M. Sundaralingam, (eds.), (Adenine Press, NY), 1988.
2032. TUNG, C. S., AND HARVEY, S. C. A molecular mechanical model to predict the helical twist angles of B-DNA. *Nucl. Acids Res.* **12** (1984), 3343-3356.
2033. TUNG, C. S., AND HARVEY, S. C. Base sequence, local helix structure, and macroscopic curvature of A-DNA and B-DNA. *J. Biol. Chem.* **261** (1986), 3700-3709.
2034. TUNG, C. -S., AND HARVEY, S. C. Computer graphics program to reveal the dependence of the gross three-dimensional structure of the B-DNA double helix on primary structure. *Nucl. Acids Res.* **14** (1986), 381-387.
2035. TUNG, C. S., AND HARVEY, S. C. A common structural feature in promoter sequences of *E. coli*. *Nucl. Acids Res.* **15** (1987), 4973-4985.
2036. TURNER, D. H., SUGIMOTO, N., AND FREIER, S. M. RNA structure prediction. *Ann. Rev. Biophys. Biophys. Chem.* **17** (1988), 167-192.
2037. TURNER, J. Approximation algorithms for the shortest common superstring problem. *Information and Computation* **83** (1989), 1-20.
2038. TYLER, E. C., HORTON, M. R., AND KRAUSE, P. R. A review of algorithms for molecular sequence comparison. *Computers and Biomedical Research* **24** (1991), 72-96.
2039. TYSON, H., AND HALEY, B. Alignment of nucleotide or amino acid sequences on microcomputers, using a modification of Sellers' (1974) algorithm which avoids the need for calculation of the complete distance matrix. *Computer Programs in Biomedicine* **21** (1985), 3-10.
2040. UCHIDA, H. DNA data bank of Japan. *Protein Nucleic Acid and Enzyme* **29** (1986), 159-162.
2041. UKKONEN, E. A linear-time algorithm for finding approximate shortest common superstrings. *Algorithmica* **5** (1990), 313-323.
2042. ULAM, S. M. Some ideas and prospects in biomathematics, (in) *Annual Review of Biophysics and Bioengineering*, Morales, M.F., (ed.), (Annual Reviews Inc., Palo Alto, CA), 1972.
2043. ULANOVSKY, L. E., AND TRIFONOV, E. N. Estimation of wedge components in curved DNA. *Nature* **326** (1987), 720-722.
2044. ULJANOV, A. V., ABAGYAN, R. A., AND SKRYABIN, K. G. MACS-effective algorithm to search signals based on physico-chemical properties of DNA, (in) *Abstracts of: Application of Data Bases and Microcomputers in Molecular Biology*, (Jena), p.37, 1988.
2045. UNGER, R., HAREL, D., AND SUSSMAN, J. L. DNAMAT: an efficient graphic matrix sequence homology algorithm and its application to structural analysis. *CABIOS* **2** (1986), 283-289.
2046. University of Texas System-Center For High Performance Computing, Management Report, 1988.
2047. U.S. Congress, Office of Technology Assessment. *Mapping Our Genes—The Genome Projects, How Big, How Fast?* OTA-BA-373 Washington, D.C.: U.S. Govt. Printing Office, 1988.
2048. USHA, R., AND MURTHY, M. R. N. Protein structural homology: a metric approach. *Int. J. Pept. Protein Res.* **28** (1986), 364-369.
2049. VAJDA, S., AND DELISI, C. Fast prediction of the backbone structure in low energy polypeptide conformations by discrete dynamic programming. *Biopolymers* **29** (1990), 1755-1772.
2050. VALIQUETTE, G., ZIMMERMAN, E. A., AND ROBERTS, J. L. mRNA sequence predictions from homologous protein sequences. *J. Theor. Biol.* **112** (1985), 445-458.
2051. VAN BOCKSTAELE, F. Sequence representation. *Biochimie* **67** (1985), 509-516.
2052. VAN DEN BERG, J. A., AND OSINGA, M. A peptide to DNA conversion program. *Nucl. Acids Res.* **14** (1986), 137-140.
2053. VAN ROMPUY, L. L., LESAGE, C., VANDERHAEGEN, M. E., TELEMAN, M. P., AND ZABEAU, M. F. An improved FORTRAN 77 recombinant DNA database management system with graphic extensions in GKS. *CABIOS* **2** (1987), 251-255.
2054. VAN THIEL, L. R., AND MICHALEWICZ, J. Restriction enzyme cleavage site identification using Microsoft Word. *Biotechniques* **5** (1987), 115.
2055. VARGA, G., CARROLL, T., HAMORI, E., AND LIM, H. A. Computer graphics algorithm for abstract representation of protein and DNA sequences using PHIGS+ and Motif program development tools, FSU-SCRI-91-150, October 1991.
2056. VASS, J. K., AND WILSON, R. H. 'ZSTATS'-A statistical analysis for potential Z-DNA surfaces. *Nucl. Acids Res.* **12** (1984), 825-832.
2057. VELJKOVICH, V., COSIC, I., DIMITRIJEVICH, B., AND LALOVIC, D. Is it possible to analyse DNA and protein sequences by the methods of digital signal processing? *IEEE Transact. Biomed. Engineer.* **BME-32** (1985), 337-341.
2058. VICKERY, L. E. Interactive analysis of protein structure using microcomputer spreadsheet. *Trends Biochem. Sci.* **12** (1987), 37-39.
2059. VIETORISZ, T. NYSERNET: Market And Economic Impact Study Volume IV: Industrial Clients, 1988.
2060. VIHINEN, M. An algorithm for simultaneous comparison of several sequences. *CABIOS* **4(1)** (1988), 89-92.

2061. VIHINEN, M. Simultaneous comparison of several sequences, (in) **Methods in Enzymology-Molecular Evolution: Computer Analysis of Protein and Nucleic Acid Sequences**, Vol.183, Doolittle, R.F., (ed.), (Academic Press, NY), pp.447-456, 1990.
2062. VINGRON, M., AND ARGOS, P. A fast and sensitive multiple sequence alignment algorithm. *CABIOS* 5(2) (1989), 115-121.
2063. VINGRON, M., AND ARGOS, P. Determination of reliable regions in protein sequence alignments. *Protein Engineering* 3(7) (1990), 565-569.
2064. VINGRON, M., AND ARGOS, P. Motif recognition and alignment for many sequences by comparison of dot-matrices. *J. Mol. Biol.* 218 (1991), 33-43.
2065. VISHKIN, U. Optimal parallel pattern matching in strings, (in) **Proceedings of the 12th ICALP**, Lecture Notes in CS 194, (Springer-Verlag, NY), pp.497-508, 1985.
2066. VISWANADHAN, V. N., DENCKLA, B., AND WEINSTEIN, J. N. New joint prediction algorithm (Q_7 -JASEP) improves the prediction of protein secondary structure. *Biochemistry* 30 (1991), 11164-11172.
2067. VOGL, G., AND ARGOS, P. Searching for distantly related protein sequences in large databases by parallel processing on a transputer machine. *CABIOS* 8(1) (1992), 49-56.
2068. VOLINIA, S., BERNARDI, F., GAMBARI, R., AND BARRAI, I. Co-localization of rare oligonucleotides and regulatory elements in mammalian upstream gene regions. *J. Mol. Biol.* 203 (1988), 385-390.
2069. VOLINIA, S., GAMBARI, R., BERNARDI, F., AND BARRAI, I. The frequency of oligonucleotides in mammalian genic regions. *Comput. Appl. Biosci.* 5 (1989), 33-40.
2070. VOLKOV, V. V., LEONTYEV, A. Yu., AND WINTER, V. G. Application of Fourier spectroscopy to characterization of origins of replication, (in) **Modelling and Computer Methods in Molecular Biology and Gentics**, Kolchanov, N.A., Ratner, V.A., (eds.), (Abstracts of the International Conference, Novosibirsk), pp.20-21, 1990.
2071. VON HAESELER, A. Limits in the reconstruction of phylogenetic trees exemplified with 5S rRNA sequences, **Conceptual and Numerical Analysis of Data**, Opitz, O., (ed.), (Spring-Verlag, NY), 1987.
2072. VON HEIJNE, G. Membrane proteins. The amino acid composition of membrane-penetrating segments. *Eur. J. Biochem.* 120 (1981), 275-278.
2073. VON HEIJNE, G. On the hydrophobic nature of signal sequences. *Eur. J. Biochem.* 116 (1981), 419-422.
2074. VON HEIJNE, G. Patterns of amino acids near signal-sequence cleavage sites. *Eur. J. Biochem.* 133 (1983), 17-21.
2075. VON HEIJNE, G. Analysis of the distribution of charged residues in the N-terminal region of signal sequences: implications for protein export in prokaryotic and eukaryotic cells. *EMBO J.* 3 (1984), 2315-2318.
2076. VON HEIJNE, G. How signal sequences maintain cleavage specificity. *J. Mol. Biol.* 173 (1984), 243-251.
2077. VON HEIJNE, G. Signal sequences. The limits of variation. *J. Mol. Biol.* 184 (1985), 99-105.
2078. VON HEIJNE, G. A new method for predicting signal sequences cleavage sites. *Nucl. Acids Res.* 14 (1986), 4683-4690.
2079. VON HEIJNE, G. **Sequence Analysis In Molecular Biology**, (Academic Press, NY), 1987.
2080. VON HEIJNE, G. SIGPEP: a sequence database for secretory signal peptides. *Protein Seq. Data Anal.* 1 (1987), 41-42.
2081. VON HEIJNE, G. Getting sense out of sequence data. *Nature* 333 (1988), 605-607.
2082. VONDERVISZT, F., MATRAI, G., AND SIMON, I. Characteristic sequential residue environment of amino acids in proteins. *Int. J. Pept. Protein Res.* 27 (1986), 483-492.
2083. VONDERVISZT, F., AND SIMON, I. A possible way for prediction of domain boundaries in globular proteins from amino acid sequence. *Biochem. Biophys. Res. Commun.* 139 (1986), 11-17.
2084. VYSOTSKAYA, G. S., GUSEV, V. D., AND KULICHKOV, V. A. Method for discovering informative zones in genetic punctuation marks, (in) **Theoretical Research and Data Banks in Molecular Biology and Genetics**, (Abstracts of the Nat. Conference, Novosibirsk), pp.54-58, 1986.
2085. WADA, A., AND SUYAMA, A. Third letters in codons counterbalance the (G+C)-content of their first and second letters. *FEBS Lett.* 188 (1985), 291-294.
2086. WADA, D. , (et al.). Codon usage tabulated from the GenBank genetic sequence data. *Nucl. Acids Res.* (1990).
2087. WAGNER, F., HART, R., FINK, R., AND CLASSEN, M. Evaluation of inter-and intramolecular primary structure homologies of interferons by a Monte Carlo method. *J. Interferon Research* 10 (1990), 31-38.
2088. WAGNER, R. A., AND FISCHER, M. J. The string-to-string correction problem. *J. ACM* 21 (1974), 168-173.
2089. WAKO, H., AND SCHERAGA, H. A. On the use of distance constraints to fold a protein. *Macromolecules* 9 (1981), 961-969.
2090. WALKER, J. R., AND WILLET, P. Compression of nucleic acid and protein sequence data. *CABIOS* 2 (1986), 89-93.
2091. WALLACE, B. A., CASCIO, M., AND MIELKE, D. L. Evaluation of methods for the prediction of membrane protein secondary structures. *Proc. Natl. Acad. Sci. USA* 83 (1986), 9423-9427.
2092. WALLACE, R. B. DNASTAR-a microcomputer-based DNA sequence management system. *Biotechnology Software* 1 (1984), 6.

2093. WANG, J. C. Interactions between twisted DNAs and enzymes: the effects of superhelical turns. *J. Mol. Biol.* **87** (1974), 797.
2094. WANG, Y. P., AND PAVLIDIS, T. Optimal correspondence of string subsequences. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **12** (1990), 1080–1087.
2095. WARME, P. K., MOMANY, F. A., RUMBALL, S. V., TUTTLE, R. V., AND SCHERAGA, H. A. Computation of structure of homologous proteins: alpha-lactalbumin from lysozyme. *Biochemistry* **13** (1974), 768–782.
2096. WATANABE, K., URANO, Y., AND TAMAOKI, T. Optimal alignments of biological sequences on a microcomputer. *CABIOS* **1** (1985), 83–87.
2097. WATANABE, K., YASUKAWA, K., AND ISO, K. Graphic display of nucleic acid structure by a microcomputer. *Nucl. Acids Res.* **12** (1984), 801–809.
2098. WATERMAN, M. D., SMITH, T. F., AND BURKS, C. The statistical distribution of nucleic acid similarities. *Nucl. Acids Res.* **13** (1985), 645–656.
2099. WATERMAN, M. S. Secondary structure of single stranded nucleic acids. *Adv. Math. Supplementary Studies* **1** (1978), 167–212.
2100. WATERMAN, M. S. Combinatorics of RNA hairpins and cloverleaves. *Stud. Appl. Math.* **60** (1979), 91–96.
2101. WATERMAN, M. S. Frequencies of restriction sites. *Nucl. Acids Res.* **24** (1983), 8951–8956.
2102. WATERMAN, M. S. Sequence alignments in the neighborhood of the optimum with general application to dynamic programming. *Proc. Natl. Acad. Sci. USA* **80** (1983), 3123–3124.
2103. WATERMAN, M. S. Efficient sequence alignment algorithms. *J. Theor. Biol.* **108** (1984), 333–337.
2104. WATERMAN, M. S. General methods of sequence comparison. *Bull. Math. Biol.* **46**(4) (1984), 473–500.
2105. WATERMAN, M. S. Dynamic programming algorithms for picture comparison. *Adv. Appl. Math.* **6** (1985), 129–134.
2106. WATERMAN, M. S. Multiple sequence alignment by concensus. *Nucl. Acids Res.* **14(22)** (1986), 9095–9102.
2107. WATERMAN, M. S. Probability distributions for DNA sequence comparisons. *Lectures in the Life Sciences* **17** (1986), 29–56.
2108. WATERMAN, M. S. Computer analysis of nucleic acid sequences, (in) **Methods in Enzymology**, Vol.164, (Academic Press, NY), pp.756–795, 1988.
2109. WATERMAN, M. S. Consensus methods for folding single-stranded nucleic acids, (in) **Mathematical Methods for DNA Sequences**, Waterman, M., (ed.), (CRC Press, Boca Raton), 1989.
2110. WATERMAN, M. S. Consensus patterns in sequences, (in) **Mathematical Methods for DNA Sequences**, Waterman, M.S., (ed.), (CRC Press, Boca Raton), pp.93-116, 1989.
2111. WATERMAN, M. S. Forward to 1989 special issue on molecular sequence analysis. *Bull. Math. Biol.* **51** (1989), 1–4.
2112. WATERMAN, M. S. Mathematical analysis of molecular sequences. *Bull. Math. Biology* **51** (1989), 1–4.
2113. WATERMAN, M. S. Sequence alignments, (in) **Mathematical Methods for DNA Sequences**, Waterman, M.S., (ed.), (CRC Press, Boca Raton), pp.53-92, 1989.
2114. WATERMAN, M. S., AND ARRATIA, R. Critical phenomena in sequence matching. *Annals of Prob.* **13** (1985), 1236–1249.
2115. WATERMAN, M. S., ARRATIA, R., AND GALAS, D. J. Pattern recognition in several sequences: Consensus and alignment. *Bull. Math. Biol.* **46** (1984), 515–527.
2116. WATERMAN, M. S., ARRATIA, R., AND GORDON, L. An extreme value theory for sequence matching. *Ann. Stat.* **14** (1986), 971–993.
2117. WATERMAN, M. S., AND BEMENT, W. A. Locating maximum variance segments in sequential data. *J. of the International Association for Mathematical Geology* **9** (1977), 55–61.
2118. WATERMAN, M. S., AND BYERS, T. H. Determining all optimal and near-optimal solutions when solving shortest path problems by dynamic programming. *Operations Research* **32** (1984), 1381–1384.
2119. WATERMAN, M. S., AND BYERS, T. H. A dynamic programming algorithm to find all solutions in a neighborhood of the optimum. *Math. Biosci.* **77** (1985), 179–188.
2120. WATERMAN, M. S., AND EGGERT, M. A new algorithm for best subsequence alignments with application to tRNA-rRNA comparisons. *J. Mol. Biol.* **197** (1987), 723–728.
2121. WATERMAN, M. S., GORDON, L., AND ARRATIA, R. Phase transitions in sequence matches and nucleic acid structure. *Proc. Natl. Acad. Sci. USA* **84** (1987), 1239–1243.
2122. WATERMAN, M. S., AND GRIGGS, J. R. Interval graphs and maps of DNA. *Bull. Math. Biol.* **48** (1986), 189–195.
2123. WATERMAN, M. S., GRIGGS, J. R., AND HANLON, P. J. Sequence alignments with matched sections. *SIAM J. Alg. Disc. Meth.* **7** (1986), 604–608.
2124. WATERMAN, M. S., AND JONES, R. Consensus methods for DNA and protein sequence alignment, (in) **Methods in Enzymology-Molecular Evolution: Computer Analysis of Protein and Nucleic Acid Sequences**, Vol.183, Doolittle, R.F., (ed.), (Academic Press, NY), pp.221-237, 1990.
2125. WATERMAN, M. S., AND PERLWITZ, M. D. Line geometries for sequence comparisons. *Bull. Math. Biol.* **46** (1984), 567–577.
2126. WATERMAN, M. S., AND SMITH, T. F. RNA secondary structure: A complete mathematical analysis. *Math. BioSci.* **42** (1978), 257–266.

2127. WATERMAN, M. S., AND SMITH, T. F. Rapid dynamic programming algorithms for RNA secondary structure. *Adv. Applied Math.* **7** (1986), 455-464.
2128. WATERMAN, M. S., SMITH, T. F., AND BEYER, W. A. Some biological sequence metrics. *Adv. Math.* **20** (1976), 367-387.
2129. WATERMAN, M. S., SMITH, T. F., AND KATCHER, H. L. Algorithms for restriction map comparisons. *Nucl. Acids Res.* **12** (1984), 237-242.
2130. WATERMAN, M. S., SMITH, T. F., SINGH, M., AND BEYER, W. A. Additive evolutionary trees. *J. Theor. Biol.* **64** (1977), 199-213.
2131. WATERMAN, M. S., AND STEIN, P. R. On some new sequences generalizing the Catalan and Motzkin numbers. *Discrete Math.* **26** (1978), 261-272.
2132. WATERMAN, M. S. , (ed.). **Mathematical Methods for DNA Sequences**, (CRC Press, Inc., Boca Raton, Florida), 1989.
2133. WATTS, S. Parallel thinking takes on the human genome. *New Scientist* **120** (1988), 34.
2134. WAYNE, M. M. Y., AND GUPTA, A. K. A Macintosh program for the management of DNA clones generated by combinatorial cassette of saturation mutagenesis. *Biotchniques* **7** (1989), 604-606.
2135. WEBER, P. C., LUKAS, T. J., CRAIG, T. A., WILSON, E., KING, M. M., KWIATKOWSKI, A. P., AND WATTERSON, D. M. Computational and site-specific mutagenesis analyses of the asymmetric charge distribution on calmodulin. *Proteins* **6** (1989), 70-85.
2136. WEINER, M. P., AND SCHERAGA, H. A. A set of Macintosh computer programs for the design and analysis of synthetic genes. *CABIOS* **5** (1989), 191-198.
2137. WEINER, P. Linear pattern matching algorithms, (in) **Conference Record, IEEE 14th Annual Symposium on Switching and Automata Theory**, pp.1-11, 1973.
2138. WEIR, B. S. Statistical analysis of molecular genetic data. *IMA J. Math. Appl. Med. Biol.* **2** (1985), 1-39.
2139. WEIR, B. S. Statistical analysis of DNA sequences. *J. Natl. Cancer Institute* **80** (1988), 495-506.
2140. WEIR, B. S. , (ed.). **Statistical Analysis of DNA Sequence Data**, (Marcel Dekker, NY), 1983.
2141. WEISE, M. J. A microcomputer program for hydrophobic analysis of proteins with I/O through word processing and graphics software. *CABIOS* **2** (1986), 103-106.
2142. WEISE, M. J. An access interface for the MS-DOS diskette format of GenBank (R), a gene sequence database. *CABIOS* **3** (1987), 313-317.
2143. WELLER, K. A microcomputer program using Dickerson's rules for comparative structural analysis of DNA sequences, (in) **Abstracts of: Application of Data Bases and Microcomputers in Molecular Biology**, (Jena), p.59, 1988.
2144. WELLER, K., AND CASTRO, E. RNA secondary structure prediction using microcomputers. **Abstracts of: Application of Data Bases and Microcomputers in Molecular Biology**, (Jena), p.35, 1988.
2145. WELLING, G. W., WEIJER, W. J., Van der Zee, R., AND WELLING-WESTER, S. Prediction of sequential antigenic regions in proteins. *FEBS Lett.* **188** (1985), 215-218.
2146. WELLS, D., BAUIS, W., AND KEDES, L. Codon usage in histone gene families of higher eukaryotes reflects functional rather than phylogenetic relationship. *J. Mol. Evol.* **23** (1986), 224-241.
2147. WEST, J. Automated sequence reading and analysis. *Nucl. Acids Res.* **16** (1988), 1847-1856.
2148. WHITE, C. T., HARDIES, S. C., HUTCHINSON, C. A., AND EDGEELL, M. H. The diagonal traverse homology search algorithm for locating similarities between two sequences. *Nucl. Acids Res.* **12(1)** (1984), 751-767.
2149. WHITE, J. H. Introduction to the geometry and topology of DNA structure, (in) **Mathematical Methods for DNA Sequences**, Waterman, M., (ed.), (CRC Press, Boca Raton), 1989.
2150. WHITE, S. H., AND JACOBS, R. E. Statistical distribution of hydrophobic residues along the length of protein chains. Implications for protein folding and evolution. *Biophysical Journal* **57** (1990), 911-921.
2151. WIERENGA, R. K., TERPSTRA, P., AND HOL, W. G. J. Prediction of the occurrence of ADP-binding beta-alpha-beta fold in proteins, using an amino acid sequence fingerprint. *J. Mol. Biol.* **187** (1986), 101-107.
2152. WILBUR, W. J., AND DAVID, J. L. Rapid similarity searches of nucleic acid and protein data banks. *Proc. Natl. Acad. Sci.* **80** (1983), 726-730.
2153. WILBUR, W. J., AND LIPMAN, D. J. The context dependent comparison of biological sequences. *SIAM J. Appl. Math.* **44** (1984), 557-567.
2154. WILBUR, W. L. Codon equilibrium I: Testing for homogenous equilibrium. *J. Mol. Evol.* **21** (1985), 168-181.
2155. WILBUR, W. L. Codon equilibrium II: Its use in estimating silent substitution rates. *J. Mol. Evol.* **21** (1985), 182-191.
2156. WILCOX, G. L., AND POLIAC, M. O. Generalization of protein structure from sequence using a large scale backpropagation network. *Proc. IJCNN II* (1989), 609.
2157. WILLIAMS, A. L., AND TINOCO, I. A dynamic programming algorithm for finding alternate RNA secondary structures. *Nucl. Acids Res.* **14** (1986), 299-315.
2158. WILLIAMS, K. M. Version 5 of the Mount-Conrad-Myers sequence analysis software package now available. *CABIOS* **4** (1988), 211.
2159. WILLIAMS, P. L., AND FITCH, W. M. Phylogeny determination using dynamically weighted parsimony method, (in) **Methods in Enzymology-Molecular Evolution: Computer Analysis of Protein and Nucleic Acid Sequences**, Vol.183, Doolittle, R.F., (ed.), (Academic Press, NY), pp.615-626, 1990.

2160. WILMOT, C. M., AND THORNTON, J. M. Analysis and prediction of the different types of beta-turns in proteins. *J. Mol. Biol.* **203** (1988), 221-232.
2161. WILSON, C. C. A procedure for analysing nucleic acid fragment geometry and stacking interactions. *Nucleosides & Nucleotides* **9**(2) (1990), 163-171.
2162. WILSON, I. A., HAFT, D. H., GETZOFF, E. D., TAINER, J. A., LERNER, R. A., AND BRENNER, S. Identical short peptide sequences in unrelated proteins can have different conformations: a testing ground for theories of immune recognition. *Proc. Natl. Acad. Sci. USA* **82** (1985), 5255-5259.
2163. WILSON, S. R. Further Statistical considerations for amino acid replacement data in protein sequences. *Mol. Biol. Evol.* **4**(6) (1987), 664-665.
2164. WINGENDER, E. Compilation of transcription regulation proteins. *Nucl. Acids Res.* **16** (1988), 1879-1902.
2165. WINKER, S., OVERBEEK, R., WOENSE, C. R., OLSEN, G. J., AND PFLUGER, N. Structure detection through automated covariance search. *CABIOS* **6**(4) (1990), 365-372.
2166. WINKLER, H. H., AND WOOD, D. O. Codon usage in selected AT-rich bacteria. *Biochimie* **70** (1988), 977-986.
2167. WITTEN, M. A discrete time model for the growth kinetics of tandem multiplicate genes as a model for the evolution of macromolecular complexity. Notices of the AMS November, 1977.
2168. WITTEN, M. Some mathematics of recombination: Evolution of complexity and genotypic modification in somatic cells-a possible model for aging and cancer effects. *Mech. Aging and Dev.* **13** (1980), 187-199.
2169. WITTEN, M. Computational biology: An overview. Institute for Mathematics and Its Applications, University of Minnesota. Reprint #586, 1989.
2170. WITTEN, M. Computational biology and medicine: An introduction, (in) *Encyclopedia of Computer Science*, Ralston, A., (ed.), in press, 1990.
2171. WITTEN, M., AND GROSS, A. Modeling DNA: Some applications of error correcting codes and information theory. Notices of the AMS, August, 1980.
2172. WITTMANN-LEIBOLD, B., AND ASHMAN, K. DZIONARA, M. On the statistical significance of homologous structures among the *Escherichia coli* ribosomal proteins. *Mol. Gen. Genet.* **196** (1984), 439-448.
2173. WOLF, H., MODROW, S., MOTZ, M., JAMESON, B. A., HERMANN, G., AND FORTSCH, B. An integrated family of amino acid sequence analysis programs. *CABIOS* **4** (1988), 187-191.
2174. WOLFENDEN, R. V., ANDERSEN, L., CULLIS, P. M., AND SOUTHGATE, C. C. F. Affinities of amino acid sides chains for solvent water. *Biochemistry* **20** (1981), 849-855.
2175. WOLFENDEN, R. V., CULLIS, P. M., AND SOUTHGATE, C. C. F. Water, protein folding, and the genetic code. *Science* **206** (1979), 575-577.
2176. WONG, A. K. C., REICHERT, T. A., COHEN, D. N., AND AYGUN, B. O. A generalized method for matching informational macromolecular code sequences. *Comput. Biol. Med.* **4** (1974), 43-57.
2177. WONG, C. K., AND CHANDRA, A. K. Bounds for the string editing problem. *J. ACM* **23** (1976), 13-16.
2178. WONG, J. T. Untitled letter. *Trends Biochem. Sci.* **6** (1981), xiii.
2179. WONG, J. T. F., AND CEDERGREN, R. Natural selection versus primitive gene structure as determinant of codon usage. *Eur. J. Biochem.* **159** (1986), 175-180.
2180. WOOD, K. GENEPRO: intelligent MS-DOS software for molecular biologists. *Biotechnology Software* **11** (1987), 10-14.
2181. WOODHEAD, A. D., AND BARNHART, B. J., (eds.). *Biotechnology and the Human Genome*, (Plenum Press, NY), 1988.
2182. WORLEY, P. H. The effects of time constraints on scaled speed-up. ORNL/TM-11031, Oak Ridge National Laboratory, 1989.
2183. WOZNIAK, P., AND MAKALOWSKI, W. Searching for tRNA genes in DNA sequences—an IBM microcomputer program. *CABIOS* **6** (1990), 49-50.
2184. WRIGHT, F. The 'effective number of codons' used in a gene. *Gene* **87** (1990), 23-29.
2185. WU, C. H., ERMONGKONCHAI, A., AND CHANG, T.-C. Protein classification using a neural network protein database (NNPDB) system, Proceedings of the 1991 Analysis of Neural Network Applications Conference, pp.29-41, 1991.
2186. WU, C. H., AND WHITSON, G. M. Neural network database systems for genetic sequence classification, Proceedings of the Eighth International Conference on Mathematical and Computer Modeling, (in press).
2187. WU, C. H., WHITSON, G. M., AND McLARTY, J. W. Artificial neural system for gene classification using a domain database. Proc. 1990 ACM 18th Annual Computer Science Conference, pp.288-292, 1990.
2188. WU, C. H., WHITSON, G. M., AND MONTLLOR, G. J. PROCANS: A protein classification system using a neural network, preprint, 1989.
2189. WU, H. M., AND CROTHERS, D. M. The locus of sequence-directed and protein-induced DNA bending. *Nature* **308** (1984), 509-513.
2190. WU, C. H. McLARTY, J. W., AND WHITSON, G. M. Neural networks for molecular sequence database management, Proceedings of the 1991 ACM 19th Annual Computer Science Conference, pp.588-594, 1991.
2191. WU, S., MANBER, U., MYERS, G., AND MILLER, W. An $O(np)$ sequence comparison algorithm. *Info. Proc. Letters* **35** (1990), 317-323.

2192. WU, T. T., FITCH, W. M., AND MARGOLIASH, E. The information content of protein amino acid sequence. *Ann. Rev. Biochem.* **43** (1974), 539–566.
2193. WU, T. T., AND KABAT, E. A. An attempt to evaluate the influence of neighboring amino acids ($n-1$) and ($n+1$) on the backbone conformation of amino acid in proteins. Use in predicting the three-dimensional structure of the polypeptide backbone of other proteins. *J. Mol. Biol.* **75** (1973), 13–31.
2194. WULKAN, M., AND LOTT, T. J. Computer-aided construction of nucleic acid restriction maps using defined vectors. *CABIOS* **1** (1985), 235–239.
2195. YAGIL, G., AND SUSSMAN, J. L. Structural models for non-helical DNA. *The EMBO Journal* **5**(7) (1986), 1719–1725.
2196. YAMAMOTO, K., KITAMURA, Y., AND YOSHIKURA, H. Computation of statistical secondary structures of nucleic acids. *Nucl. Acids Res.* **12** (1984), 335–346.
2197. YAMAMOTO, K., SAKURAI, N., AND YOSHIKURA, H. Graphics of RNA secondary structure: towards an object-oriented algorithm. *CABIOS* **3**(2) (1987), 99–103.
2198. YAMAMOTO, K., AND YOSHIKURA, H. Computer program for prediction of the optimal and suboptimal secondary structures of long RNA molecules. *CABIOS* **1** (1985), 89–94.
2199. YAMAMOTO, K., AND YOSHIKURA, H. A new representation of protein structure: vector diagram. *CABIOS* **2** (1986), 83–88.
2200. YAMAMOTO, K., AND YOSHIKURA, H. Relation between genomic and capsid structures in RNA viruses. *Nucl. Acids Res.* **14** (1986), 389–396.
2201. YAMAMOTO, K., AND YOSHIKURA, H. An improved algorithm for the prediction of optimum and suboptimum folding structures of long single-stranded RNA. *CABIOS* **3** (1987), 31–35.
2202. YANG, J., YE, J., AND WALLACE, D. C. Computer selection of oligonucleotide probes from amino acid sequences for use in the gene library screening. *Nucl. Acids Res.* **12** (1984), 837–843.
2203. YANG, J. T., WU, C. -S. C., AND MARTINEZ, H. M. Calculation of protein conformation from circular dichroism, (in) *Methods in Enzymology*, Vol.130, (Academic Press, NY), pp.208–269, 1986.
2204. YANOFSKY, C. Attenuation in the control of expression of bacterial operons. *Nature* **289** (1981), 751–758.
2205. YAO, F. F. Efficient dynamic programming using quadrangle inequality, (in) *Proceedings of the ACM Symposium on Theory of Computing*, pp.429–435, 1980.
2206. YARUS, M., AND FOLLEY, L. S. Sense codons are found in specific context. *J. Mol. Biol.* **182** (1985), 529–540.
2207. YASUOKA, N., AND IGA, Y. Computer graphics representation of protein structure. *Protein, Nucleic Acid and Enzyme* **29** (1986), 136–142.
2208. YEH, L. S., HUNT, L. T., GEORGE, D. G., AND BARKER, W. C. Searching protein sequences for potential transmembrane segments. *Fed. Proc.* **46** (1987), 2232.
2209. YOCKEY, H. P. A prescription which predicts functionally equivalent residues at given sites in protein sequences. *J. Theor. Biol.* **67** (1977), 337–343.
2210. YOMO, T., AND OHNO, S. Concordant evolution of coding and noncoding regions of DNA made possible by the universal rule of TA/CG deficiency-TG/CT excess. *Proc. Natl. Acad. Sci. USA* **86** (1989), 8452–8456.
2211. YOUNG, F. W. Structured data analysis and data analysis strategies. August 17, 1989. preprint.
2212. ZAHN, K., AND BLATTNER, F. R. Sequence-induced DNA curvature at the bacteriophage lambda origin of replication. *Nature* **317** (1985), 451–453.
2213. ZEHETNER, G., FRISCHAUF, A., AND LEHRACH, H. Approaches to restriction map determination, (in) *Nucleic Acid and Protein Sequence Analysis a Practical Approach*, Bishop, M.J., Rawlings, C.J., (eds.), (IRL Press, Oxford), pp.359–385, 1987.
2214. ZERIAL, M., SALINAS, J., FILIPSKI, J., AND BERNARDI, G. Gene distribution and nucleotide sequence organization in the human genome. *Eur. J. Biochem.* **160** (1986), 479–485.
2215. ZHANG, C. -T. Equations between frequencies of amino acids in organisms. *J. Theor. Biol.* **142** (1990), 281–284.
2216. ZHANG, K., AND SHASHA, D. Simple fast algorithms for the editing distance between trees and related problems. *SIAM J. Computing* **18** (1989), 1245–1262.
2217. ZHURKIN, V. B. Periodicity in DNA primary structure is defined by secondary structure of the coded protein. *Nucl. Acids Res.* **9** (1981), 1963–1971.
2218. ZHURKIN, V. B., ULYANOV, N. B., GORIN, A. A., AND JERNIGAN, R. L. Static and statistical bending of DNA evaluated by Monte Carlo simulations. *Proc. Natl. Acad. Sci. USA* **88** (1991), 7046–7050.
2219. ZHURKIN, V. B., ULYANOV, N. B., AND IVANOV, V. I. Mechanisms of DNA bending in the free state and in the nucleosome, (in) *Structure and Expression-DNA Bending and Curvature*, Vol.3., Olson, W.K., Sarma, M.K., Sarma, R.H., Sundaralingam, M., (eds.), (Adenine Press, NY), p.169, 1988.
2220. ZIENTARA, G. P., AND NAGY, J. A. Proteins and polypeptides: computer graphics for space-filling model representations. *Computers and Chemistry* **7** (1983), 67–74.
2221. ZIMMERMAN, J. M., ELIEZER, N., AND SIMHA, R. The characterization of amino acid sequences in proteins by statistical methods. *J. Theor. Biol.* **21** (1968), 170–201.
2222. ZUKER, M. RNA folding prediction: The continued need for interaction between biologists and mathematicians, (in) *Lectures on Mathematics in The Life Sciences: Some Mathematical Questions in Biology DNA Sequence Analysis*, Miura, R., (ed.), (American Mathematical Society, Providence, RI), 1986.

2223. ZUKER, M. On finding all suboptimal foldings of an RNA molecule. *Science* **244** (1989), 48–52.
2224. ZUKER, M. The use of dynamic programming algorithms in RNA secondary structure prediction, (in) **Mathematical Methods for DNA Sequences**, Waterman, M., (ed.), (CRC Press, Boca Raton), 1989.
2225. ZUKER, M., AND SANKOFF, D. RNA secondary structures and their prediction. *Bull. Math. Biol.* **46** (1984), 591–621.
2226. ZUKER, M., AND SOMORJAI, R. L. The alignment of protein structures in 3 dimensions. *Bull. Math. Biol.* **51**(1) (1989), 55–78.
2227. ZUKER, M., AND STIEGLER, P. Optimal computer folding of large RNA sequences using thermodynamics and auxiliary information. *Nucl. Acids Res.* **9**(1) (1981), 133–148.
2228. ZULL, J. E., AND SMITH, S. K. Is genetic code redundancy related to retention of structural information on both DNA strands? *Trends Biochem. Sci.* **15** (1990), 257–261.
2229. ZVELEBIL, M. J., BARTON, G. J., TAYLOR, W. R., AND STERNBERG, M. J. E. Prediction of protein secondary structure and active sites using the alignment of homologous sequences. *J. Mol. Biol.* **195** (1987), 957–961.
2230. ZWEIG, S. E. Analysis of large nucleic acid dot matrices on small computers. *Nucl. Acids Res.* **12** (1984), 767–776.
2231. ZWERLING, H., MEHLORN, R., PACKER, L., AND McELROY. A computer technique for structural studies of bacteriorhodopsin, (in) **Methods in Enzymology**, Vol.88, (Academic Press, NY), pp.772-784, 1982.