References

Aho, A. V., R. Sethi, et al. (1985). Compilers: Principles, Techniques, and Tools, Addison-Wesley.

Alizadeh, A. A., M. B. Eisen, et al. (2000). "Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling." Nature **403**(6769): 503-11.

Altschul, S. F., W. Gish, et al. (1990). "Basic local alignment search tool." J Mol Biol **215**(3): 403-10.

Asai, K., S. Hayamizu, et al. (1993). "Prediction of protein secondary structure by the hidden Markov model." Comput Appl Biosci **9**(2): 141-6.

Bailey, T. L. and C. Elkan (1994). "Fitting a mixture model by expectation maximization to discover motifs in biopolymers." Proc Int Conf Intell Syst Mol Biol **2**: 28-36.

Bairoch, A. (2000). "The ENZYME database in 2000." Nucleic Acids Res **28**(1): 304-5.

Bateman, A., E. Birney, et al. (2000). "The Pfam protein families database." Nucleic Acids Res **28**(1): 263-6.

Bates, M. D., C. R. Erwin, et al. (2002). "Novel genes and functional relationships in the adult mouse gastrointestinal tract identified by microarray analysis." Gastroenterology **122**(5): 1467-82.

Baum, L. E., T. Petrie, et al. (1970). "A maximization technique occuring in the statistical analysis of probabilitic functions of Markov chains." The Annals of Mathematical Statistics **41**(1): 164171.

Benson, D. A., I. Karsch-Mizrachi, et al. (2003). "GenBank." Nucleic Acids Res **31**(1): 23-7.

Birney, E. and R. Durbin (1997). "Dynamite: a flexible code generating language for dynamic programming methods used in sequence comparison." Proc Int Conf Intell Syst Mol Biol **5**: 56-64.

Birney, E. and R. Durbin (2000). "Using GeneWise in the Drosophila annotation experiment." Genome Res **10**(4): 547-8.

Bishop, C. M. and T. M. E. (2000). Variational relevance vector machines. Proceedings of the 16th Conference on Uncertainty in Artificial Intelligence, Morgan Kaufmann.

Boeckmann, B., A. Bairoch, et al. (2003). "The SWISS-PROT protein knowledgebase and its supplement TrEMBL in 2003." Nucleic Acids Res **31**(1): 365-70.

Bowman, S., D. Lawson, et al. (1999). "The complete nucleotide sequence of chromosome 3 of Plasmodium falciparum." Nature **400**(6744): 532-8.

Brenton, J. D., S. A. Aparicio, et al. (2001). "Molecular profiling of breast cancer: portraits but not physiognomy." Breast Cancer Res **3**(2): 77-80.

Brown, M., R. Hughey, et al. (1993). "Using Dirichlet mixture priors to derive hidden Markov models for protein families." Proc Int Conf Intell Syst Mol Biol **1**: 47-55.

Buck, K. J., R. A. Harris, et al. (1991). "A general method for quantitative PCR analysis of mRNA levels for members of gene families: application to GABAA receptor subunits." Biotechniques **11**(5): 636-41.

Burge, C. and S. Karlin (1997). "Prediction of complete gene structures in human genomic DNA." J Mol Biol **268**(1): 78-94.

Butte, A. J. and I. S. Kohane (2000). "Mutual information relevance networks: functional genomic clustering using pairwise entropy measurements." Pac Symp Biocomput: 418-29.

Butte, A. J., J. Ye, et al. (2001). "Determining significant fold differences in gene expression analysis." Pac Symp Biocomput: 6-17.

Chen, T., H. L. He, et al. (1999). "Modeling gene expression with differential equations." Pac Symp Biocomput: 29-40.

Chu, S., J. DeRisi, et al. (1998). "The transcriptional program of sporulation in budding yeast." Science 282(5389): 699-705.

Churchill, G. A. (1989). "Stochastic models for heterogeneous DNA sequences." Bull Math Biol 51(1): 79-94.

Corcoran, L. M., J. K. Thompson, et al. (1988). "Homologous recombination within subtelomeric repeat sequences generates chromosome size polymorphisms in P. falciparum." Cell 53(5): 807-13.

Dawson, E., G. R. Abecasis, et al. (2002). "A first-generation linkage disequilibrium map of human chromosome 22." Nature 418(6897): 544-8.

Dib, C., S. Faure, et al. (1996). "A comprehensive genetic map of the human genome based on 5,264 microsatellites." Nature 380(6570): 152-4.

Dixon, D. A. and S. C. Kowalczykowski (1991). "Homologous pairing in vitro stimulated by the recombination hotspot, Chi." Cell 66(2): 361-71.

Dowell, R. D., R. M. Jokerst, et al. (2001). "The Distributed Annotation System." BMC Bioinformatics 2(1): 7.

Down, T. (2003). Genetics. Cambridge, Cambridge.

Dunham, I., N. Shimizu, et al. (1999). "The DNA sequence of human chromosome 22." Nature 402(6761): 489-95.

Durbin, R., Eddy, E., Krogh A. and Mitchison, G. (1998). Biological Seqence Analysis. Cambridge, UK, Cambridge University Press.

Eddy, S. R. (2001). "Profile hidden markov models for biological sequence analysis." Bioinformatics 14: 755-763.

Eisen, M. B., P. T. Spellman, et al. (1998). "Cluster analysis and display of genome-wide expression patterns." Proc Natl Acad Sci U S A 95(25): 14863-8.

Escalante, A. A., A. A. Lal, et al. (1998). "Genetic polymorphism and natural selection in the malaria parasite Plasmodium falciparum." Genetics 149(1): 189-202.

Falquet, L., M. Pagni, et al. (2002). "The PROSITE database, its status in 2002." Nucleic Acids Res 30(1): 235-8.

Figueiredo, L. M., L. H. Freitas-Junior, et al. (2002). "A central role for Plasmodium falciparum subtelomeric regions in spatial positioning and telomere length regulation." Embo J 21(4): 815-24.

Gamma, E., R. Helm, et al. (1994). Design Patterns: Elements of Reusable Object-Oriented Software, Addison Wesley Professional.

Gardner, M. J., N. Hall, et al. (2002). "Genome sequence of the human malaria parasite Plasmodium falciparum." Nature 419(6906): 498-511.

Gardner, M. J., H. Tettelin, et al. (1999). "The malaria genome sequencing project: complete sequence of Plasmodium falciparum chromosome 2." Parassitologia 41(1-3): 69-75.

Gendrel, C. G., A. Boulet, et al. (2000). "(CA/GT)(n) microsatellites affect homologous recombination during yeast meiosis." Genes Dev 14(10): 1261-8.

Gilbert, S. F. (2003). Developmental Biology. Sunderland, MA, Sinauer Associates, Inc.

Gosling, J., B. Joy, et al. (2000). The Java Language Specification, Addison-Wesley.

Grandjean, F., L. Bremaud, et al. (2001). "Sequential gene expression of P-glycoprotein (P-gp), multidrug resistance-associated protein (MRP) and lung resistance protein: functional activity of P-gp and MRP present in the doxorubicin-resistant human K562 cell lines." Anticancer Drugs **12**(3): 247-58.

Grundy, W. N., T. L. Bailey, et al. (1997). "Meta-MEME: motif-based hidden Markov models of protein families." Comput Appl Biosci **13**(4): 397-406.

Guillouzic, S., I. I. L'Heureux, et al. (2000). "Rate processes in a delayed, stochastically driven, and overdamped system." Phys Rev E Stat Phys Plasmas Fluids Relat Interdiscip Topics **61**(5A): 4906-14.

Guo, Z., R. A. Guilfoyle, et al. (1994). "Direct fluorescence analysis of genetic polymorphisms by hybridization with oligonucleotide arrays on glass supports." Nucleic Acids Res **22**(24): 5456-65.

Hall, N., A. Pain, et al. (2002). "Sequence of Plasmodium falciparum chromosomes 1, 3-9 and 13." Nature **419**(6906): 527-31.

Hasan, S. (2003). The Sanger Centre. Cambridge.

Huang, A., D. Fuchs, et al. (2002). "Serum tryptophan decrease correlates with immune activation and impaired quality of life in colorectal cancer." Br J Cancer **86**(11): 1691-6.

Hubbard, T., D. Barker, et al. (2002). "The Ensembl genome database project." Nucleic Acids Res **30**(1): 38-41.

Hughes, T. R., M. Mao, et al. (2001). "Expression profiling using microarrays fabricated by an ink-jet oligonucleotide synthesizer." Nat Biotechnol **19**(4): 342-7.

Hughey, R. and A. Krogh (1995). SAM: Sequence alignment and modeling software system. Santa Cruz, CA, University of California.

Iwagaki, H., A. Hizuta, et al. (1995). "Decreased serum tryptophan in patients with cancer cachexia correlates with increased serum neopterin." Immunol Invest **24**(3): 467-78.

Kihara, C., T. Tsunoda, et al. (2001). "Prediction of sensitivity of esophageal tumors to adjuvant chemotherapy by cDNA microarray analysis of gene-expression profiles." Cancer Res **61**(17): 6474-9.

Krogh, A., M. Brown, et al. (1994). "Hidden Markov models in computational biology. Applications to protein modeling." J Mol Biol **235**(5): 1501-31.

Lashkari, D. A., J. L. DeRisi, et al. (1997). "Yeast microarrays for genome wide parallel genetic and gene expression analysis." Proc Natl Acad Sci U S A **94**(24): 13057-62.

Lawrence, T. S. (1988). "Reduction of doxorubicin cytotoxicity by ouabain: correlation with topoisomerase-induced DNA strand breakage in human and hamster cells." Cancer Res **48**(3): 725-30.

Lawrence, T. S. and M. A. Davis (1990). "The influence of Na+,K(+)-pump blockade on doxorubicin-mediated cytotoxicity and DNA strand breakage in human tumor cells." Cancer Chemother Pharmacol **26**(3): 163-7.

Li, Y., C. Campbell, et al. (2002). "Bayesian automatic relevance determination algorithms for classifying gene expression data." Bioinformatics **18**(10): 1332-9.

Liefers, G. J. and R. A. Tollenaar (2002). "Cancer genetics and their application to individualised medicine." Eur J Cancer **38**(7): 872-9.

Lindholm, T. and F. Yellin (1999). The Java Virtual Machine Specification, Addison-Wesley.

Majewski, J. and J. Ott (2000). "GT repeats are associated with recombination on human chromosome 22." Genome Res **10**(8): 1108-14.

Mody, M., Y. Cao, et al. (2001). "Genome-wide gene expression profiles of the developing mouse hippocampus." Proc Natl Acad Sci U S A **98**(15): 8862-7.

Nedelman, J., P. Heagerty, et al. (1992). "Quantitative PCR with internal controls." Comput Appl Biosci **8**(1): 65-70.

Nelder, J. A. and P. McCulagh (1983). Generalized Linear Models. London, Chapman and Hall.

Ning, Z., A. J. Cox, et al. (2001). "SSAHA: a fast search method for large DNA databases." Genome Res **11**(10): 1725-9.

O'Hagan, A. (1994). Bayesian Inference. London, Hodder Anold.

Perou, C. M., T. Sorlie, et al. (2000). "Molecular portraits of human breast tumours." Nature **406**(6797): 747-52.

Platt, J. (1998). Fast Training of Support Vector Machines using Sequential Minimal Optimization. Advances in Kernel Methods - Support Vector Learning. B. C. a. S. Scholkopf B., A., MIT Press.

Pocock, M. R., T. Down, et al. (2000). "BioJava: Open Source Components for Bioinformatics." sigbio newsletter **20**(2): 10-12.

Pollack, Y., A. L. Katzen, et al. (1982). "The genome of Plasmodium falciparum. I: DNA base composition." Nucleic Acids Res **10**(2): 539-46.

Potmesil, M., Y. H. Hsiang, et al. (1988). "Resistance of human leukemic and normal lymphocytes to drug-induced DNA cleavage and low levels of DNA topoisomerase II." Cancer Res **48**(12): 3537-43.

Pourquier, P., D. Montaudon, et al. (1998). "Doxorubicin-induced alterations of c-myc and c-jun gene expression in rat glioblastoma cells: role of c-jun in drug resistance and cell death." Biochem Pharmacol **55**(12): 1963-71.

Rabiner, L. R. (1989). "A tutorial on Hidden Markov Models and selected applications in speech recognition." Proceedings of the IEEE **77**(2): 257-286.

Ramaswamy, S., P. Tamayo, et al. (2001). "Multiclass cancer diagnosis using tumor gene expression signatures." Proc Natl Acad Sci U S A **98**(26): 15149-54.

Rampone, S. (1998). "Recognition of splice junctions on DNA sequences by BRAIN learning algorithm." Bioinformatics **14**(8): 676-84.

Reinhardt, A. and T. Hubbard (1998). "Using neural networks for prediction of the subcellular location of proteins." Nucleic Acids Res **26**(9): 2230-6.

Rice, P., I. Longden, et al. (2000). "EMBOSS: the European Molecular Biology Open Software Suite." Trends Genet **16**(6): 276-7.

Rooney, D. E. (2001). Human Cytogenetics: Constitutional Analysis, A Practical Approach. Oxford, Oxford University Press.

Rost, B. and C. Sander (1994). "Combining evolutionary information and neural networks to predict protein secondary structure." Proteins **19**(1): 55-72.

Schena, M., D. Shalon, et al. (1995). "Quantitative monitoring of gene expression patterns with a complementary DNA microarray." Science **270**(5235): 467-70.

Shalon, D. (1998). "Gene expression micro-arrays: a new tool for genomic research." Pathol Biol (Paris) **46**(2): 107-9.

Sjolander, K., K. Karplus, et al. (1996). "Dirichlet mixtures: a method for improved detection of weak but significant protein sequence homology." Comput Appl Biosci **12**(4): 327-45.

Smith, T. F. and M. S. Waterman (1981). "Identification of common molecular subsequences." <u>J Mol Biol</u> **147**(1): 195-7.

Smolen, P., D. A. Baxter, et al. (2001). "Modeling circadian oscillations with interlocking positive and negative feedback loops." <u>J Neurosci</u> **21**(17): 6644-56.

Stajich, J. E., D. Block, et al. (2002). "The Bioperl toolkit: Perl modules for the life sciences." <u>Genome Res</u> **12**(10): 1611-8.

Stoesser, G., W. Baker, et al. (2003). "The EMBL Nucleotide Sequence Database: major new developments." <u>Nucleic Acids Res</u> **31**(1): 17-22.

Tewey, K. M., T. C. Rowe, et al. (1984). "Adriamycin-induced DNA damage mediated by mammalian DNA topoisomerase II." <u>Science</u> **226**(4673): 466-8.

Tipping, M. E. (2000). <u>The Relevance Vector Machine</u>. Advances in Neural Information Processing Systems 12, MIT Press.

van 't Veer, L. J., H. Dai, et al. (2002). "Gene expression profiling predicts clinical outcome of breast cancer." <u>Nature</u> **415**(6871): 530-6.

Vapnik, V. N. (1995). The Nature of Statistical Learning Theory, Springer.

Womble, D. D. (2000). "GCG: The Wisconsin Package of sequence analysis programs." <u>Methods Mol Biol</u> **132**: 3-22.