

# Bibliography

This bibliography lists both works on specific topics whose authors are mentioned in the text and sources consulted and felt to be valuable for further reading, but not specifically cited; their utility should be apparent from the title of the book or article.

- C. Adami & N.J. Cerf, Physical complexity of symbolic sequences. *Physica D* 137 (2000) 62–69.
- M. Ageno, Linee di ricerca in fisica biologica. *Accad. Naz. Lincei* 102 (1967) 3–50.
- M. Ageno, *La “Macchina” Batterica*. Rome: Lombardo Editore (1992).
- S.E. Ahnert, T.M.A. Fink & A. Zinovyev, How much non-coding DNA do eukaryotes require? *J. theor. Biol.* 252 (2008) 587–592.
- R. Albert & A.-L. Barabási, Statistical mechanics of complex networks. *Rev. mod. Phys.* 71 (2002) 47–97.
- P.M. Allen, Evolving complexity in social science. In: *Systems—New Paradigms for the Human Sciences* (G. Altman & W.A. Koch, eds). Berlin: Walter de Gruyter (1998).
- L. Allison, D. Powell & T.I. Dix, Compression and approximate matching. *Computer J.* 42 (1999) 1–10.
- W. Arber, Molecular mechanisms of biological evolution. In: *Frontiers in Biology* (C.-H. Chou & K.-T. Shao, eds), pp. 19–24. Taipei: Academia Sinica (1998).
- R.B. Ash, *Information Theory*. New York: Interscience (1965).
- R.B. Ash, *A Primer of Abstract Mathematics*. Washington, DC: Mathematical Association of America (1998).
- W.R. Ashby, *An Introduction to Cybernetics*. London: Chapman & Hall (1956).
- W.R. Ashby, Principles of the self-organizing system. In: *Principles of Self Organization* (H. von Foerster & G.W. Zopf, eds), pp. 255–278. Oxford: Pergamon Press (1962).
- B. Audit, N. Audit, C. Vaillant et al., Long-range correlations between DNA bending sites: relation to the structure and dynamics of nucleosomes. *J. mol. Biol.* 316 (2002) 903–918.
- H. Baharvand, Embryonic stem cells: Establishment, maintenance and differentiation. In: *Embryonic Stem Cell Research* (E.V. Grier, ed.), pp. 1–63. Hauppauge, NY: Nova Science (2006).
- P. Bak & K. Sneppen, Punctuated equilibrium and criticality in a simple model of evolution. *Phys. Rev. Lett.* 71 (1993) 4083–4086.
- P.G. Baker, C.A. Goble, S. Bechhofer et al., An ontology for bioinformatics applications. *Bioinformatics* 15 (1999) 510–520.
- E.N. Baker & R.E. Hubbard, Hydrogen bonding in globular proteins. *Prog. Biophys. mol. Biol.* 44 (1984) 97–179.
- W. Banzhaf, G. Beslon, S. Christensen et al., From artificial evolution to computational evolution. *Nature Rev. Genet.* 7 (2006) 729–735.
- A.D. Baxevanis & B.F.F. Ouellette (eds), *Bioinformatics*, 2nd edn. New York: Wiley (2001).
- D. Benedetto, E. Caglioti & V. Loreto, Language trees and zipping. *Phys. Rev. Lett.* 88 (2002) 048702.

- C.H. Bennett, Logical depth and physical complexity. In: *Universal Turing Machine—A Half-Century Survey* (R. Herken, ed.), pp. 227–257. Oxford: University Press (1988).
- E. Bernstein & C.D. Allis, RNA meets chromatin. *Genes Dev.* 19 (2005) 1635–1655.
- L. von Bertalanffy, *Théorie Générale des Systèmes*. Paris: Dunod (1993).
- M. Bibikova, Zh. Lin, L. Zhou et al., High-throughput DNA methylation profiling using universal bead arrays. *Genome Res.* 16 (2006) 383–393.
- P. Billingsley, Statistical methods in Markov chains. *Ann. math. Statist.* 32 (1961) 12–40.
- G.M. Blackburn & M.J. Gait, *Nucleic Acids in Chemistry and Biology*, 2nd edn, pp. 210–221. Oxford: University Press (1996).
- L.A. Blumenfeld, *Problems of Biological Physics*. Berlin: Springer (1981).
- B. Bollobás, *Graph Theory*. New York: Springer (1979).
- P. Borwein & L. Jörgenson, Visible structures in number theory. *Am. math. Monthly* 108 (2001) 897–910.
- I. Braslavsky, B. Hebert, E. Kartalov & S.R. Quake, Sequence information can be obtained from single DNA molecules. *Proc. Natl Acad. Sci. USA* 100 (2003) 3960–3964.
- R.F. Bruinsma, Physics of protein-DNA interaction. *Physica A* 313 (2002) 211–237.
- M.G. Cacace, E.M. Landau & J.J. Ramsden, The Hofmeister series: Salt and solvent effects on interfacial phenomena. *Q. Rev. Biophys.* 30 (1997) 241–278.
- L.L. Campbell, Entropy as a measure. *IEEE Trans. Inform. Theory* IT-11 (1965) 112–114.
- R. Carnap & Y. Bar-Hillel, *An Outline of a Theory of Semantic Information*. MIT Research Laboratory of Electronics Technical Report No 247 (1952).
- J. Chelly, J.P. Concordet, J.C. Kaplan & A. Kahn, Illegitimate transcription: transcription of any gene in any cell type. *Proc. Natl Acad. Sci. USA* 86 (1989) 2617–2621.
- D.S. Chernavsky, Synergetics and information. *Matematika Kibernetika* 5 (1990) 3–42 (in Russian).
- C. Cherry, *On Human Communication*. London: Chapman & Hall (1957).
- S. Chumakov, C. Belapurkar, C. Putonti et al., The theoretical basis of universal identification systems for bacteria and viruses. *J. biol. Phys. Chem.* 5 (2005) 121–128.
- T. Coenye & P. Vandamme, Use of the genomic signature in bacterial classification and identification. *System. Appl. Microbiol.* 27 (2004) 175–185.
- J. Costello & C. Plass, Methylation matters. *J. med. Genet.* 38 (2001) 285–303.
- F.H.C. Crick, L. Barnett, S. Brenner & R.J. Watts-Tobin, General nature of the genetic code for proteins. *Nature* 192 (1961) 1227–1232.
- J.P. Crutchfield, The calculi of emergence. *Physica D* 75 (1994) 11–54.
- F. Despa, A. Fernández & R.S. Berry, Dielectric modulation of biological water. *Phys. Rev. Lett.* 93 (2004) 228104.
- T.G. Dewey, Algorithmic complexity of a protein. *Phys. Rev. E* 54 (1996) R39–R41.
- T.G. Dewey, Algorithmic complexity and thermodynamics of sequence-structure relationships in proteins. *Phys. Rev. E* 56 (1997) 4545–4552.
- P. Duesberg, R. Li, A. Fabarius & R. Hehlmann, The chromosomal basis of cancer. *Cell. Oncol.* 27 (2005) 293–318.
- R.A. Dwek & T.D. Butters (eds), *Glycobiology*. *Chem. Rev.* 102 (2002) no 2 (pp. 283 ff.).
- A.W.F. Edwards, *Likelihood*. Cambridge: University Press (1972).
- M. Eigen, Wie entsteht Information? *Ber. Bunsenges.* 76 (1976) 1059–1081.
- R.J. Ellis, Macromolecular crowding. *Trends Biochem. Sci.* 26 (2001) 597–604.
- P. Érdi & Gy. Barna, Self-organizing mechanism for the formation of ordered neural mappings. *Biol. Cybernet.* 51 (1984) 93–101.
- D.A. Fell, Metabolic control analysis: A survey of its theoretical and experimental development. *Biochem. J.* 286 (1992) 313–330.
- W. Feller, *An Introduction to Probability Theory and its Applications*, 3rd edn, Vol. 1. New York: Wiley (1967).
- G. Felsenfeld & M. Groudine, Controlling the double helix. *Nature* 421 (2003) 448–453.
- A. Fernández, Pause sites and regulatory role of secondary structure in RNA replication. *Biophys. Chem.* 34 (1989) 29–33.

- A. Fernández, Correlation of pause sites in MDV-1 RNA replication with kinetic refolding of the growing chain. A Monte Carlo simulation of the Markov process. *Eur. J. Biochem.* 182 (1989) 161–163.
- A. Fernández & H. Cendra, In vitro RNA folding: The principle of sequential minimization of entropy loss at work. *Biophys. Chem.* 58 (1996) 335–339.
- A. Fernández & R. Scott, Dehydron: A structurally encoded signal for protein interaction. *Biophys. J.* 85 (2003) 1914–1928.
- A. Fernández, T.R. Sosnick & A. Colubri, Dynamics of hydrogen bond desolvation in protein folding. *J. mol. Biol.* 321 (2002) 659–675.
- A. Fernández, J. Kardos, L.R. Scott et al., Structural defects and the diagnosis of amyloidogenic propensity. *Proc. Natl Acad. Sci. USA* 100 (2003) 6446–6451.
- J.W. Fickett, The gene identification problem: An overview for developers. *Computers Chem.* 20 (1996) 103–118.
- R.A. Fisher, *The Design of Experiments*, 6th edn. Edinburgh: Oliver & Boyd (1951).
- S.P.A. Fodor, J.L. Read, M.C. Pirrung et al., Light-directed, spatially addressable parallel chemical synthesis. *Science* 251 (1991) 767–773.
- H. von Foerster, On self-organizing systems and their environments. In: *Self-Organizing Systems* (M.C. Yorvitz & S. Cameron, eds). Oxford: Pergamon (1960).
- L.T.C. França, E. Carrilho & T.B. Kist, A review of DNA sequencing techniques. *Q. Rev. Biophys.* 35 (2002) 169–200.
- H. Frauenfelder, From atoms to biomolecules. *Helv. Phys. Acta* 57 (1984) 165–187.
- S. Galam & A. Mauger, Universal formulas for percoation thresholds. *Phys. Rev. E* 53 (1996) 2177–2181; *ibid.* 55 (1997) 1230–1231.
- M.R. Gardner & W.R. Ashby, Connectance of large dynamic (cybernetic) systems: Critical values for stability. *Nature* 228 (1970) 784.
- M. Gell-Mann & S. Lloyd, Information measures, effective complexity, and total information. *Complexity* 2 (1996) 44–52.
- C. Gibas & P. Jambeck, *Developing Bioinformatics Computer Skills*. Sebastopol, Cal.: O'Reilly & Associates (2001).
- I.J. Good, Botryological speculations. In: *The Scientist Speculates* (I.J. Good, ed.), pp. 120–132. London: Heinemann (1962).
- I.J. Good, Statistics of language. In: *Encyclopaedia of Linguistics, Information and Control* (A.R. Meetham, ed.), pp. 567–581. Oxford: Pergamon Press (1969).
- A.N. Gorban, T.G. Popova & M.G. Sadovsky, Classification of symbol sequences over their frequency dictionaries. *Open Syst. Inform. Dynam.* 7 (2000) 1–17.
- A.N. Gorban, T.G. Popova & A. Zinovyev, Codon usage trajectories and 7-cluster structure of 143 complete bacterial genomic sequences. *Physica A* 353 (2005) 365–387.
- S.J. Gould, *Ontogeny and Phylogeny*. Cambridge, Mass.: Belknap Press (1977).
- A. Graps, An introduction to wavelets. *IEEE Comput. Sci. Engng* 2 (1995) 50–61.
- P. Grassberger, Toward a quantitative theory of self-generated complexity. *Int. J. theor. Phys.* 25 (1986) 907–938.
- B.F. Gray, Reversibility and biological machines. *Nature* 253 (1975) 436–437; *ibid.* 257 (1975) 72.
- S.F. Gull & G.J. Daniell, Image reconstruction from incomplete and noisy data. *Nature* 272 (1978) 686–690.
- S.P. Gygi, B. Rist, S.A. Gerber et al., Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nature Biotechnol.* 17 (1999) 994–999.
- R.W. Hamming, Error detecting and error correcting codes. *Bell Syst. tech. J.* 26 (1950) 147–160.
- W.P. Hanage, C. Fraser & B.G. Spratt, Sequences, sequence clusters and bacterial species. *Phil. Trans. R. Soc. B* 361 (2006) 1917–1927.
- R.V.L. Hartley, Transmission of information. *Bell Syst. tech. J.* 7 (1928) 535–563.
- B. Hartmann & R. Lavery, DNA structural forms. *Q. Rev. Biophys.* 29 (1996) 309–368.
- H. Hillman, *The Case for New Paradigms in Cell Biology and in Neurobiology*. Lewiston: Edwin Mellen Press (1991).
- R. Hooke, *Micrographia*. London: The Royal Society (1665).

- B.A. Huberman & T. Hogg, Complexity and adaptation. *Physica D* 22 (1986) 376–384.
- R. Jaenisch, DNA methylation and imprinting. *Trends Genet.* 13 (1997) 323–329.
- M. Jakobsson, S.W. Scholz, P. Scheet et al., Genotype, haplotype and copy-number variation in worldwide human populations. *Nature* 451 (2008) 998–1003.
- P. James, Protein identification in the post-genome era. *Q. Rev. Biophys* 30 (1997) 279–331.
- T. Jenuwein & C.D. Allis, Translating the histone code. *Science* 293 (2001) 1074–1080.
- H. Jeong, S.P. Mason, A.-L. Barabási & Z.N. Oltvai, Lethality and centrality in protein networks. *Nature* 411 (2001) 41–42.
- S.A. Johnson & T. Hunter, Kinomics: Methods for deciphering the kinome. *Nature Methods* 2 (2005) 17–25.
- T.B. Jongeling, Self-organization and competition in evolution: A conceptual problem in the use of fitness landscapes. *J. theor. Biol.* 178 (1996) 369–373.
- S. Karlin & V. Brendel, Patchiness and correlations in DNA sequences. *Science* 259 (1993) 677–679.
- S. Karlin, I. Ladunga & B.E. Blaisdell, Heterogeneity of genomes: Measures and values. *Proc. Natl Acad. Sci. USA* 91 (1994) 12837–12841.
- S.A. Kauffman, Emergent properties in random complex automata. *Physica D* 10 (1984) 145–156.
- E. Kellenberger, Assembly in biological systems. *Ciba Foundation Symp. (New Ser.)* 7 (1972) 189–206.
- E.S. Kempner & J.H. Miller, The molecular biology of *Euglena gracilis* IV. Cellular stratification by centrifuging. *Expl Cell Res.* 51 (1968) 141–149; *idem*, V. Enzyme localization. *Expl Cell Res.* 51 (1968) 150–156.
- D.G. Kendall, A mathematical approach to seriation. *Phil. Trans. R. Soc. A* 269 (1970) 125–135.
- F. Képès & C. Vaillant, Transcription-based solenoidal model of chromosomes. *Complexus* 1 (2003) 171–180.
- P.M. Kim, L.J. Lu, Y. Xia & M.B. Gerstein, Relating three-dimensional structures to protein networks provides evolutionary insights. *Science* 314 (2006) 1938–1941.
- W. Kolch, M. Calder & D. Gilbert, When kinases meet mathematics. *FEBS Lett.* 579 (2005) 1891–1895.
- A.N. Kolmogorov, Three approaches to the quantitative definition of information. *Probl. Peredachi Inform.* 1 (1965) 3–11.
- A.N. Kolmogorov, Combinatorial foundations of information theory and the calculus of probabilities. *Usp. Mat. Nauk.* 38 (1983) 27–36.
- A.N. Kolmogorov & V.A. Uspenskii, Algorithms and randomness. *Theor. Prob. Appl.* 32 (1988) 389–412.
- A.A. Kornyshev & S. Leikin, Sequence recognition in the pairing of DNA duplexes. *Phys. Rev. Lett.* 86 (2001) 3666–3669.
- J.B. Kruskal, Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrika* 29 (1964) 1–27.
- S. Kullback & R.A. Leibler, On information and sufficiency. *Ann. math. Statist.* 22 (1951) 79–86.
- A. Lempel & J. Ziv, On the complexity of finite sequences. *IEEE Trans. Info. Theory* IT-22 (1976) 75–81.
- V.G. Levich, *Physicochemical Hydrodynamics*. Englewood Cliffs, NJ: Prentice-Hall (1962).
- J.C. Lindon, J.K. Nicholson, E. Holmes & J.R. Everett, Metabonomics: Metabolic processes studied by NMR spectroscopy of biofluids. *Concepts magn. Reson.* 12 (2000) 289–320.
- H. Lipson, Principles of modularity, regularity, and hierarchy for scalable systems. *J. biol. Phys. Chem.* 7 (2007) 125–128.
- P.O. Luthi, A. Preiss, B. Chopard & J.J. Ramsden, A cellular automaton model for neurogenesis in *Drosophila*. *Physica D* 118 (1998) 151–160.
- S. Lloyd & H. Pagels, Complexity as thermodynamic depth. *Ann. Phys.* 188 (1988) 186–213.
- B. McClintock, The origin and behavior of mutable loci in maize. *Proc. Natl Acad. Sci. USA* 36 (1950) 344–355.

- E.H. McConkey, Molecular evolution, intracellular organization, and the quinary structure of proteins. *Proc. Natl Acad. Sci. USA* 79 (1982) 3236–3240.
- D.M. Mackay, Quantal aspects of scientific information. *Phil. Mag.* (ser. 7) 41 (1950) 289–311.
- D.M. Mackay, Operational aspects of intellect. In: *Mechanization of Thought Processes*, NPL Symposium No 10, pp. 37–73. London: HMSO (1960).
- J.D. Madura, M.E. Davis, M.K. Gilson et al., Biological applications of electrostatic calculations and Brownian dynamics simulations. *Rev. comput. Chem.* 5 (1994) 229–267.
- B. Mandelbrot, Contribution à la théorie mathématique des jeux de communication. *Publ. Inst. Statist. Univ. Paris* 2 (1952) 1–124.
- S. Manghani & J.J. Ramsden, The efficiency of chemical detectors. *J. biol. Phys. Chem.* 3 (2003) 11–17.
- A.A. Markov, Statistical analysis of the text of “Eugene Onegin” illustrating the connexion with investigations into chains. *Izv. Imp. Akad. Nauk.*, Ser. 6, No 3 (1913) 153–162 (in Russian).
- S. Martin, Zh. Zhang, A. Martino & J.-L. Faulon, Boolean dynamics of genetic regulatory networks inferred from microarray time series data. *Bioinformatics* 23 (2007) 866–874.
- A.M. Mood, The distribution theory of runs. *Ann. math. Statist.* 11 (1940) 367–392.
- Y.-H. Oh, Y.-P. Kim & H.-S. Kim, SUMO chip for analysis of SUMO-conjugation to a target protein. *Biochip J.* 1 (2007) 28–34.
- P.D. Patel & G. Weber, Electrophoresis in free fluid: A review of technology and agrifood applications. *J. biol. Phys. Chem.* 3 (2003) 60–73.
- L. Peliti, Fitness landscapes and evolution. In: *Physics of Biomaterials* (T. Riste & D. Sherrington, eds), pp. 287–308. Dordrecht: Kluwer (1996).
- J.K. Percus, O.E. Percus & A.S. Perelson, Predicting the size of the T-cell receptor and antibody combining region from consideration of efficient self-nonsel self discrimination. *Proc. Natl Acad. Sci. USA* 90 (1993) 1691–1695.
- M. Planck, The concept of causality. *Proc. Phys. Soc.* 44 (1932) 529–539.
- G.H. Pollack, *Cells, Gels and the Engines of Life*. Seattle: Ebner (2001).
- O. Popescu & G.N. Misevic, Self-recognition by proteoglycans. *Nature* 386 (1997) 231–232.
- D.J. Raine & V.J. Norris, Network structure of metabolic pathways. *J. biol. Phys. Chem.* 1 (2002) 89–94.
- J.J. Ramsden, The photolysis of small silver halide particles. *Proc. R. Soc. Lond. A* 392 (1984) 427–444.
- J.J. Ramsden, Computing photographic response curves. *Proc. R. Soc. Lond. A* 406 (1986) 27–37.
- J.J. Ramsden, Experimental methods for investigating protein adsorption kinetics at surfaces. *Q. Rev. Biophys.* 27 (1994) 41–105.
- J.J. Ramsden, Kinetics of protein adsorption. In: *Biopolymers at Interfaces* (M. Malmsten, ed.), pp. 321–361. New York: Dekker (1998).
- J.J. Ramsden, The specificity of biomolecular particle adhesion. *Colloids Surf. A* 173 (2000) 237–249.
- J.J. Ramsden & J. Dreier, Kinetics of the interaction between DNA and the type IC restriction enzyme *EcoR124/3I*. *Biochemistry* 35 (1996) 3746–3753.
- J.J. Ramsden & M. Grätzel, Formation and decay of methyl viologen radical cation dimers on the surface of colloidal CdS. *Chem. phys. Lett.* 132 (1986) 269–272.
- J.J. Ramsden & J. Vohradský, Zipf-like behavior in procaryotic protein expression. *Phys. Rev. E* 58 (1998) 7777–7780.
- J.J. Ramsden, D.J. Roush, D.S. Gill et al., Protein adsorption kinetics drastically altered by repositioning a single charge. *J. Am. Chem. Soc.* 117 (1995) 8511–8516.
- A. Rényi, Kémiai reakciók tárgyalása a sztochasztikus folyamatok elmélete segítségével. *Magy. Tud. Akad. Mat. Kut. Int. Közl.* 2 (1953) 83–101.
- A. Rényi, *Probability Theory*. Budapest: Akadémiai Kiadó (1970).
- E.J. Richards & S.C.R. Elgin, Epigenetic codes for heterochromatin formation and silencing. *Cell* 108 (2002) 489–500.

- J.C. Robinson, All possible chaotic dynamics can be approximated in three dimensions. *Nonlinearity* 11 (1998) 529–545.
- J.M. Romanovsky, N.V. Stepanova & D.S. Chernavsky, *Kinetische Modelle in der Biophysik*. Jena: Gustav Fischer (1974).
- V. Ye. Ruzhentsev, The problem of transition in palaeontology. *Int. Geol. Rev.* 6 (1964) 2204–2213.
- F. Sanger, Determination of nucleotide sequences in DNA. *Biosci. Rep.* 1 (1981) 3–18.
- G. Sarkar & S.S. Sommer, Access to a messenger RNA sequence or its protein product is not limited by tissue or species specificity. *Science* 244 (1989) 331–334.
- M.A. Savageau, Comparison of classical and autogenous systems of regulation in inducible operons. *Nature* 252 (1974) 546–549.
- F. Scheller & F. Schubert, *Biosensoren*. Berlin: Akademie-Verlag (1989).
- K. Scherrer & J. Jost, The gene and the genom concept. *Mol. syst. Biol.* 3:87 (2007) (pp. 1–11).
- P. Schuck, Kinetics of ligand binding to receptors immobilized in a polymer matrix, as detected with an evanescent wave biosensor. I. A computer simulation of the influence of mass transport. *Biophys. J.* 70 (1996) 1230–1249.
- C.E. Shannon, A mathematical theory of communication. *Bell Syst. tech. J.* 27 (1948) 379–423.
- C.E. Shannon, Prediction and entropy of printed English. *Bell Syst. tech. J.* 30 (1951) 50–64.
- C.E. Shannon & W. Weaver, *The Mathematical Theory of Communication*. Urbana: University of Illinois Press (1949).
- R. Shaw, Strange attractors, chaotic behaviour, and information flow. *Z. Naturforsch.* 36a (1981) 80–112.
- A.R. Sheldrake, The ageing, growth and death of cells. *Nature* 215 (1974) 381–385.
- O. Sinanoğlu, What size cluster is like a surface? *Chem. Phys. Lett.* 81 (1981) 188–190.
- P. Smolen, D.A. Baxter & J.H. Byrne, Mathematical modeling of gene networks. *Neuron* 26 (2000) 567–580.
- A.K. Solomon, Red cell membrane structure and ion transport. *J. gen. Physiol.* 43 (1960) 1–15.
- G. Sommerhoff, *Analytical Biology*. London: Oxford University Press (1950).
- S.M. Stanley, A theory of evolution above the species level. *Proc. Natl Acad. Sci. USA* 72 (1975) 646–650.
- S.C. Stearns, The evolutionary significance of phenotypic plasticity. *BioScience* 39 (1989) 436–445.
- G. Stent, Explicit and implicit semantic content of the genetic information. In: *The Centrality of Science and Absolute Values*, 4th Int. Conf. on the Unity of the Sciences, Vol. 1, pp. 261–277. New York: International Cultural Foundation (1975).
- M.C.R. Symons, Water structure and reactivity. *Acc. chem. Res.* 14 (1981) 179–187.
- R.J. Taft, M. Pheasant & J.S. Mattick, The relationship between non-protein-coding DNA and eukaryotic complexity. *BioEssays* 29 (2007) 288–299.
- T.M. Thompson, *From Error-Correcting Codes through Sphere Packings to Simple Groups*. Washington, DC: Mathematical Association of America (1983).
- N. Tkemaladze, On the problems of an automated system of pattern recognition with learning. *J. biol. Phys. Chem.* 2 (2002) 80–84.
- G.M. Tomkins, The metabolic code. *Science* 189 (1975) 760–763.
- S. Tonegawa, Somatic generation of antibody diversity. *Nature* 302 (1983) 575–581.
- H.G. Trüper, How to name a prokaryote? *FEMS Microbiol. Rev.* 23 (1999) 231–249.
- R. Tureck, Cells, functions, relationships in musical structure and performance. *Proc. R. Inst.* 67 (1995) 277–318.
- J.-M. Valletton, Information processing in biomolecule-based biomimetic systems. *Reactive Polymers* 12 (1990) 109–131.
- R.A. VanBogelen, K.D. Greis, R.M. Blumenthal et al., Mapping regulatory networks in microbial cells. *Trends Microbiol.* 7 (1999) 320–327.
- S.B. Volchan, What is a random sequence? *Am. math. Monthly* 109 (2002) 46–63.
- J. Vohradský, Neural network model of gene expression. *FASEB J.* 15 (2001) 846–854.

- J. Vohradský & J.J. Ramsden, Genome resource utilization during procaryotic development. *FASEB J.* 15 (2001) 2054–2056.
- R.F. Voss, Evolution of long-range fractal correlations and  $1/f$  noise in DNA base sequences. *Phys. Rev. Lett.* 68 (1992) 3805–3808.
- T.T. Vu & J. Vohradský, Nonlinear differential equation model for quantification of transcriptional regulation applied to microarray data of *Saccharomyces cerevisiae*. *Nucl. Acids Res.* 35 (2007) 279–287.
- B.L. van der Waerden, Beweis einer Baudet'schen Vermutung. *Nieuw. Arch. Wiskunde* 15 (1927) 212–216.
- A. Wagner & D.A. Fell, The small world inside large metabolic networks. *Proc. R. Soc. Lond. B* 268 (2001) 1803–1810.
- J.D. Watson & F.H.C. Crick, Molecular structure of nucleic acids. *Nature* 171 (1953) 737–738.
- D.J. Watts & S.H. Strogatz, Collective dynamics of 'small-world' networks. *Nature* 393 (1998) 440–442.
- V. Welby, Significs. *Encyclopaedia Britannica* (11th edn). (1911).
- N. Wiener, *Cybernetics, or Control and Communication in the Animal and the Machine* (Actualités Sci. Ind. no 1053). Paris: Hermann & Cie. (1948).
- S.H. White, Global statistics of protein sequences. *A. Rev. Biophys. biomol. Structure* 23 (1994) 407–439.
- C.R. Woese, G.J. Olsen, M. Ibba & D. Soll, Aminoacyl-tRNA synthetases, the genetic code, and the evolutionary process. *Microbiol. mol. Biol. Rev.* 64 (2000) 202–236.
- S. Wolfram, Statistical mechanics of cellular automata. *Rev. mod. Phys.* 55 (1983) 601–644.
- S. Wright, Character change, speciation and the higher taxa. *Evolution* 36 (1982) 427–443.
- A.A. Zamyatin, A.S. Borchikov, M.G. Vladimirov & O.L. Voronina, The EROP-Moscow oligopeptide database. *Nucleic Acids Res.* 34 (2006) D261–D266.
- J. Ziv & A. Lempel, A universal algorithm for sequential data compression. *IEEE Trans. Inform. Theory* IT-23 (1977) 337–343.
- H.L. Zaaijer, F.J. van Hemert, M.H. Koppelman & V.V. Lukashov, Independent evolution of overlapping polymerase and surface protein genes of hepatitis B virus. *J. gen. Virol.* 88 (2007) 2137–2143.
- W. H. Zurek, Thermodynamic cost of computation, algorithmic complexity, and the information metric. *Nature* 341 (1989) 119–124.

# Index

## A

accelerated network, 91  
accuracy, 23  
adaptation, 157  
additive processes, 64  
algorithmic complexity, 69, 79, 92  
algorithmic compression, 70  
algorithmic information content, 79  
algorithmic information distance, 81  
allele, 130  
amino acid, 173  
aneuploidy, 130  
apoptosis, 142  
autocorrelation function, 77  
automatic annotation, 193  
automaton, 74, 86

## B

Bayes's theorem, 58, 65  
Bernoulli trials, 59  
Bernoulli, D., 22, 68  
bibliome, 257  
bilayer, 180  
biosensor, 234, 236  
bits, 11  
BLAST, 197  
blockiness, 138  
Boltzmann, L., 14  
Boolean automata, 86  
Boolean network, 89, 92, 224  
Born repulsion, 228  
Bose-Einstein statistics, 50  
brain, 4  
Brownian motion, 78

## C

C-value paradox, 136  
cancer, 130  
capacity, 33

causality, 53  
cell, 123  
cell cycle, 129  
cell membrane, 124  
cell structure, 125  
cellular automata, 88  
central dogma, 1, 123, 222  
channel, 31  
chemical genomics, 4  
chemogenomics, 4  
chromatin, 139, 148  
chromatin immunoprecipitation, 233  
chromatography, 213, 235  
chromosome, 129–131, 148  
chromosome structure, 139  
cladistics, 203  
classification, 202  
cliquishness, 90  
clustering, 103, 203, 209  
clustering coefficient, 90  
coding, 32, 35  
coenetic variable, 117  
combination, 49  
comparative genomics, 3  
complement, 47  
complexity of copies, 80  
computational biology, 4  
computational complexity, 82  
computational proteomics, 187  
conditional algorithmic information, 81  
conditional complexity, 81  
conditional information, 15  
conditional probability, 57  
connectivity, 86  
consensus sequence, 198  
constraint, 13, 74  
context, 25  
contig, 192  
control point, 249



cooperative binding, 230  
 correlated expression, 233  
 correlation information, 21, 80  
 crosslinking, 233  
 crossover, 144  
 cytoplasm, 123

## D

Darwin, C., 155  
 database reliability, 194  
 decoding, 32  
 dehydron, 176, 227, 229  
 density information, 21  
 depth, 80  
 developing embryo, 25  
 development, 146  
 differential entropy, 14  
 differentiation, 152  
 diffusion, 76, 93  
 digital organism, 4  
 dimensional reduction, 230  
 diploidy, 130  
 direct affinity measurement, 235  
 directive correlation, 118, 119  
 disorder, 14  
 distance metrics, 105  
 diversity, 11  
 DNA structure, 167  
 donor-acceptor interaction, 228  
 durability of information, 22  
 dynamic chaos, 73, 96  
 dynamic programming, 196

## E

edge complexity, 92  
 Edman sequencing, 214  
 effect, 27  
 effective complexity, 80  
 effective measure complexity, 21, 80  
 electron acceptor, 165  
 electron donor, 165  
 electrophoresis, 214  
 electrostatic interaction, 228  
 elementary flux mode, 243  
 entelechy, 151  
 entropy, 13, 14  
 entropy of a Markov process, 75  
 entropy of the source, 15  
 epigenesis, 153  
 epigenetics, 25, 148, 151, 187  
 equivocation, 43  
 ergodicity, 41, 72  
 error detection, 45  
 error rate threshold, 143, 159

eukaryote, 125  
 event, 53, 55  
 evolution, 177  
 evolution, models, 158  
 exaptation, 157  
 exon, 133, 160, 187  
 exon shuffling, 134  
 expectation, 62  
 explicit meaning, 25, 115  
 exponential growth, 85, 128  
 expressed sequence tags, 192

## F

Förster resonance, 234  
 FASTA, 197  
 feedback, 83  
 Fermi-Dirac statistics, 50  
 Fisher, R.A., 15, 83  
 focal condition, 117  
 forensic medicine, 247  
 frequency dictionary, 200  
 frequentist concept, 53  
 function, 30  
 functional cloning, 246, 249  
 functional genomics, 3, 186, 189  
 fuzzy clustering, 209

## G

G-value paradox, 136  
 gel electrophoresis, 212  
 gene, 133  
 gene expression profile, 222, 233, 249  
 gene structure, 133  
 generalized union, 56  
 genetic algorithm, 100  
 genetic code, 36  
 genetic linkage, 130  
 genome, 133  
 genome structure, 133  
 genome variation, 160  
 genon, 133  
 geological eras, 161  
 gradualism, 157  
 graph, 89

## H

Hamming distance, 24, 159, 195, 198, 202  
 haploidy, 130  
 haplotype, 138, 246  
 Hardy-Weinberg rule, 130  
 Hartley index, 12  
 heliograph, 34  
 heterogeneity, 138  
 hidden Markov model, 74, 122, 199, 218, 245

hierarchicality, 122  
 hierarchy, 10, 92  
 higher-order Markov process, 18  
 histone, 139, 148  
 Holliday junction, 144  
 homeotic genes, 155  
 homologous recombination, 143  
 homology, 185, 194  
 hybridization, 206  
 hydrogen bond, 166, 228  
 hydrophobic effect, 229  
 hypergeometric distribution, 63  
 hypotheses, 15, 65

**I**

immune repertoire, 140  
 immune system, 26, 143  
 implicit meaning, 25, 115  
 imprinting, 132  
 incompressibility, 69  
 information generation, 15, 96  
 information reception, 96  
 information science, 9  
 information theory, 9  
 instability, 86  
 integration, 119  
 interactome, 222  
 intergenomic sequence, 134  
 intersection, 47, 55  
 intron, 133  
 inverse probability, 65

**J**

joint algorithmic complexity, 81

**K**

kinetic mass action law, 227  
 Kolmogorov complexity, 72, 79  
 Kolmogorov information, 17  
 Kullback-Leibler distance, 20

**L**

lateral gene transfer, 204  
 Lewis acid, 165  
 Lewis acid-base interaction, 228  
 Lewis base, 165  
 life, 123  
 Lifshitz-van der Waals force, 228  
 likelihood ratio, 66  
 linear discriminant analysis, 105  
 linguistics, 200  
 logical depth, 82  
 logistic equation, 85  
 logon, 17  
 Lyapunov number, 94

**M**

machine with input, 121  
 Mandelbrot coding, 39  
 Markov chain, 86, 199  
 Markov chain Monte Carlo, 74  
 Markov process, 18  
 Markovian machine, 121  
 mass action law, 226  
 mass spectrometry, 214  
 maximum entropy, 68  
 meaning, 24, 82  
 meiosis, 132  
 mellitin, 173  
 memory, 60, 70, 119  
 memory function, 229, 237  
 Mendel's laws, 130  
 metabolic code, 37, 243  
 metabolic control analysis, 242  
 metabolism, 239  
 metabolite, 239  
 methylation, 132, 148  
 metron, 17  
 Michal, G., 243  
 microarray, 206, 234, 237  
 microRNA, 149  
 mismatch, 196, 230  
 missing information, 16  
 mitosis, 131  
 modularity, 122  
 module, 227  
 motif, 198  
 mRNA, 187  
 mRNA processing, 149  
 MudPIT, 214  
 multinomial coefficient, 50  
 multiplication rule, 48  
 multiplicative processes, 64  
 mutation, 145  
 mutual algorithmic complexity, 80  
 mutual algorithmic information, 81

**N**

natural selection, 115, 156  
 negative binomial distribution, 61  
 network complexity, 92  
 network diameter, 90  
 neurophysiology, 4  
 nonspecific interaction, 230  
 nucleic acid extraction, 190  
 nucleotide frequencies, 200

**O**

observations, 16  
 operon, 133, 137, 147, 223, 224

optical microscopy, 125  
 organism, 115  
 origin of proteins, 160  
 OWLS, 236

## P

parameter, 121  
 partitioning, 50  
 patchiness, 138  
 pattern discovery, 102  
 pattern recognition, 241  
 percolation, 88  
 permutation, 48  
 persistence length, 169  
 phage display, 216  
 pharmacogenomics, 250  
 phase portrait, 95  
 phosphorylation, 32  
 physical information, 17  
 physical structure, 87  
 Poisson approximation, 60  
 Poisson distribution, 57, 60  
 polymerase chain reaction, 191  
 polyploidy, 130  
 posttranslational modifications, 206  
 power spectrum, 77  
 pragmatics, 29  
 primary structure, 179  
 principal component analysis, 105, 210, 241  
 Principle of Insufficient Reason, 53, 65  
 probabilistic Boolean network, 224  
 production complexity, 80  
 prokaryote, 125  
 promoter, 135  
 promoter sites, 149  
 proofreading, 142  
 proposition, 65  
 protease, 211  
 proteasome, 130, 164  
 protein, 187  
 protein chips, 237  
 protein degradation, 130  
 protein folding, 175  
 protein interaction, 175  
 protein structure, 179  
 protein structure determination, 178  
 punctuated equilibrium, 157  
 purely random process, 73  
 pyrosequencing, 192

## Q

quality of information, 23  
 quartz crystal microbalance, 235  
 quasispecies, 143

quaternary structure, 179  
 quinary structure, 179, 222

## R

Rényi, A., 90  
 Ramachandran plot, 174  
 random graph, 90  
 random variable, 61  
 random walk, 230  
 randomness, 18, 19, 48  
 reaction-diffusion equation, 93  
 recombination, 132, 143  
 redundancy, 20, 45  
 regularity, 3, 19, 65, 70, 79, 80, 122  
 relative entropy, 20  
 remembering, 14  
 repair, 142  
 repetition, 135  
 repetitive DNA, 134  
 replication, 141  
 reptation, 170  
 response, 117  
 restriction enzymes, 207  
 retrotransposon, 139  
 ribosome, 150, 256  
 RNA folding, 172  
 RNA interference, 149  
 runs, 62

## S

sample space, 53, 55  
 sampling, 48, 51  
 Sanger, F., 191  
 satellite, 135  
 scale-free network, 91, 244  
 scatter matrix, 62  
 secondary structure, 179  
 secretome, 124  
 selection, 154, 156  
 SELEX, 234  
 self-organization, 156  
 semantic information, 26  
 semantics, 257  
 semiotics, 30, 32  
 sequence alignment, 194  
 sequence comparison, 194  
 sequencing, 191  
 serial analysis of gene expression, 210  
 Shannon coding, 39  
 Shannon index, 12  
 Shannon-Wiener index, 12  
 short interspersed element, 134  
 shotgun sequencing, 192  
 signal, 198

signalling cascades, 218  
significs, 28  
simplicity, 79  
Simpson's index, 11  
single-nucleotide polymorphism, 138, 246  
small systems, 127  
small world, 90, 244  
specificity, 230  
spontaneous assembly, 126  
standard deviation, 62  
Stark degradation, 214  
state structure, 86, 222  
stem cells, 152  
Stirling's formula, 49  
stochastic independence, 59  
strange attractor, 96  
structural complexity, 92  
structural genomics, 186  
support, 65  
surface plasmon resonance, 236  
survival, 69  
sustained activation, 231  
sustained interaction, 227  
synergetics, 242  
syntax, 19  
system, 83, 242  
systems biology, 225  
systems theory, 242

**T**

taxonomy, 161  
telomere, 130, 135  
tertiary structure, 179  
text mining, 5  
thermodynamic depth, 80  
toxicogenomics, 250

transcription, 146  
transcription factors, 149  
transcription regulation, 146  
transducer, 32, 121  
translation, 150  
tree, 91  
Turing machine, 72  
two-dimensional gel electrophoresis, 211, 215

**U**

ubiquitin, 130  
ultrastructure, 164  
unconditional complexity, 79  
unconditional information, 15, 66  
union, 47, 55

**V**

vaccine, testing of a, 61  
value of information, 22  
variance, 62  
variety, 11  
vocabulary, 187

**W**

wavelets, 106  
weighting scientific data, 22  
white noise, 77  
Whittle's formula, 74  
wrapping, 176

**Y**

yeast two-hybrid assay, 232

**Z**

Zipf's law, 40