

BIBLIOGRAPHY

- Abdelhak, S., Kalatzis, V., Heilig, R., Compain, S., Samson, D., Vincent, C., Weil, D., Cruaud, C., Sahly, I., Leibovici, M., Bitner-Glindzicz, M., Francis, M., Lacombe, D., Vigneron, J., Charachon, R., Boven, K., Bedbeder, P., Van Regemorter, N., Weissenbach, J. and Petit, C. 1997. A human homologue of the *Drosophila* eyes absent gene underlies branchio-otorenal (BOR) syndrome and identifies a novel gene family. *Nature Genetics* 15:157-164.
- Abita, J. P., Delaage, M. and Lazdunski, M. 1969. The mechanism of activation of trypsinogen. The role of the four N- terminal aspartyl residues. *European Journal of Biochemistry* 8:314-324.
- Allan, B. J., Zager, N. I., and Keller, P. J. 1970. Human pancreatic proteins: amylase, proelastase, and trypsinogen. *Archives of Biochemistry and Biophysics* 136:529-540.
- Allison, D. P., Kerper, P. S., Doktycz, M. J., Thundat, T., Modrich, P., Larimer, F. W., Johnson, D. K., Hoyt, P. R., Mucenski, M. L., and Warmack, R. J. 1997. Mapping individual cosmid DNAs by direct AFM imaging. *Genomics* 41:379-384.
- Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman 1990. Basic local alignment search tool. *Journal of Molecular Biology* 215:403-410.
- Anantharaman, T. S., Mishra, B., and Schwartz, D. 1997. Genomics via optical mapping II: ordered restriction maps. *Journal of Computational Biology* 4(2):91-118.
- Ásgeirsson, B., Fox, J. W., and Bjarnason, J. B. 1989. Purification and characterization of trypsin from the poikilotherm *Gadus morhua*. *European Journal of Biochemistry* 180:85-94.
- Barrett, A. J., and Rawlings, N. D. 1993. The many evolutionary lines of peptidases. In "Innovations in Proteases and Their Inhibitors" (F. X. Avilés, Ed.), pp. 13-30, Walter De Gruyter, Berlin.
- Bartunik, H. D., Summers, L. J. and Bartsch, H. H. 1989. Crystal structure of bovine beta-trypsin at 1.5 Å resolution in a crystal form with low molecular packing density. Active site geometry, ion pairs and solvent structure. *Journal of Molecular Biology* 210:813-828.

- Baticle, E. 1935. Le problème de la répartition. *Comptes Rendus* 201:862-864.
- Bier, M., and Nord, F. F. 1951. On the mechanism of enzyme action. XLVI. The effect of certain ions on crystalline trypsin and reinvestigation of its isoelectric point. *Archives of Biochemistry and Biophysics* 33:320-332.
- Bishop, D. T., Cannings, C., Skolnick, M., and Williamson, J. A. 1983. The number of polymorphic DNA clones required to map the human genome. In "Statistical Analysis of DNA Sequence Data" (Weir, B. S., Ed.), pp. 181-200, Marcel Dekker, New York.
- Blattner, F. R., Plunkett, G., Bloch, C. A., Perna, N. T., Burland, V., Riley, M., Collado-Vides, J., Glasner, J. D., Rode, C. K., Mayhew, G. F., Gregor, J., Davis, N. W., Kirkpatrick, H. A., Goeden, M. A., Rose, D. J., Mau, B., and Shao, Y. 1997. The complete genome sequence of *Escherichia coli* K-12. *Science* 277:1453-1474.
- Bode, W. and Schwager, P. 1975. The single calcium-binding site of crystallin bovin beta-trypsin. *FEBS Letters* 56:139-143.
- Bodenteich, A., S. Chissoe, Y.-F. Wang, and B. A. Roe. 1994. Shotgun Cloning as the Method of Choice to Generate Templates for High-throughput Dideoxynucleotide Sequencing. In "Automated DNA Sequencing and Analysis" (Adams, M., Fields, C., and Venter, J., Eds.), pp. 42-50, Academic Press, New York.
- Bordo, D., Djinovi'c, K. and Bolognesi, M. 1994. Conserved patterns in the Cu,Zn superoxide dismutase family. *Journal of Molecular Biology* 238:366-386.
- Borts, R. H. and Haber, J. E. 1987. Meiotic recombination in yeast: alteration by multiple heterozygosities. *Science* 237:1459-1465.
- Bradshaw, R. A., Neurath, H., Tye, R. W., Walsh, K. A., and Winter, W. P. 1970. Comparison of the partial amino-acid sequence of dogfish trypsinogen with bovine trypsinogen. *Nature* 226:237-239.
- Brenner, S. 1988. The molecular evolution of genes and proteins: a tale of two serines. *Nature* 334:528-530.
- Bricteux-Grégoire, S. 1970. Mise en évidence des deux trypsinogènes chez le chevreuil.

Archives Internationales de Physiologie et de Biochimie 78:399-413.

Bricteux-Grégoire, S., Schyns, R., and Florkin, M. 1971. Phylogeny of activation peptides of trypsinogen. In "Biochemical Evolution and the Origin of Life" (Schoffeniels, E., Ed.), pp. 130-149, North Holland Publishing Company, Amsterdam.

Bricteux-Grégoire, S., Schyns, R., and Florkin, M. 1971. Purification, propriétés et séquence N-terminale du trypsinogène de chèvre. *Biochimica et Biophysica Acta* 229:123-135.

Bricteux-Grégoire, S., Schyns, R., and Florkin, M. 1972. Phylogeny of trypsinogen activation peptides. *Comparative Biochemistry and Physiology* 42B:23-39.

Bricteux-Grégoire, S., Schyns, R., and Florkin, M. 1974. N-terminal amino acid sequence of trypsinogen from the elephant seal *Mirounga leonina* L. (*Carnivora*). *Biochimica et Biophysica Acta* 351:87-91.

Bricteux-Grégoire, S., Schyns, R., Florkin, M., Emmens, M., Welling, G. W., and Beintema, J. J. 1975. N-terminal amino acid sequence of trypsinogen from the lesser rorqual *Balaenoptera acutorostrata* (Cetacea). Simultaneous isolation of trypsinogen, chymotrypsinogen and ribonuclease from pancreas. *Biochimica et Biophysica Acta* 386:244-255.

Brown, D. D., Wensink, P. C. and Jordan, E. 1972. A comparison of the ribosomal DNA's of *Xenopus laevis* and *Xenopus mulleri*: the evolution of tandem genes. *Journal of Molecular Biology* 63:57-73.

Buck, F. F., Vithayathil, A. J., Bier, M., and Nord, F. F. 1962. On the mechanism of enzyme action. LXXIII. Studies on trypsin from beef, sheep, and pig pancreas. *Archives of Biochemistry and Biophysics* 97:417-424.

Bult, C. J., White, O., Olsen, G. J., Zhou, L., Fleischmann, R. D., Sutton, G. G., Blake, J. A., FitzGerald, L. M., Clayton, R. A., Gocayne, J. D., Kerlavage, A. R., Dougherty, B. A., Tomb, J. F., Adams, M. D., Reich, C. I., Overbeek, R., Kirkness, E. F., Weinstock, K. G., Merrick, J. M., Glodek, A., Scott, J. L., Geoghagen, N. S. M. and Venter, J. C. 1996. Complete genome sequence of the methanogenic archaeon, *Methanococcus jannaschii*. *Science* 273:1058-1073.

- Burland, V., Daniels, D. L., Plunkett, G., and Blattner, F. R. 1993. Genome sequencing on both strands: the Janus strategy. *Nucleic Acids Research* 21(15):3385-3390.
- Bystroff, C., and Baker, D. 1998. Prediction of local structure in proteins using a library of sequence-structure motifs. *Journal of Molecular Biology* 281(3):565-77.
- Carne, T., and Scheele, G. 1982. Amino acid sequences of transport peptides associated with canine exocrine pancreatic proteins. *Journal of Biological Chemistry* 257(8):4133-4140.
- Carreira, S., Fueri, C., Chaix, J. C. and Puigserver, A. 1996. Dietary modulation of the mRNA stability of trypsin isozymes and the two forms of secretory trypsin inhibitor in the rat pancreas. *European Journal of Biochemistry* 239:117-123.
- Charles, M., Rovey, M., Guidoni, A., and Desnuelle, P. 1963. Sur le trypsinogène et la trypsine de porc. *Biochimica et Biophysica Acta* 69:115-129.
- Chen, C. N., Su, Y., Baybayan, P., Siruno, A., Nagaraja, R., Mazzarella, R., Schlessinger, D., and Chen, E. 1996. Ordered shotgun sequencing of a 135 kb Xq25 YAC containing ANT2 and four possible genes, including three confirmed by EST matches. *Nucleic Acids Research* 24:4034-4041.
- Chen, E. Y., Schlessinger, D., and Kere, J. 1993. Ordered shotgun sequencing, a strategy for integrating mapping and sequencing of YAC clones. *Genomics* 17:651-656.
- Chen, L., DeVries, A. L. and Cheng, C. H. 1997a. Convergent evolution of antifreeze glycoproteins in Antarctic notothenioid fish and Arctic cod. *Proceedings of the National Academy of Sciences* 94:3817-3822.
- Chen, L., DeVries, A. L. and Cheng, C. H. 1997b. Evolution of antifreeze glycoprotein gene from a trypsinogen gene in Antarctic notothenioid fish. *Proceedings of the National Academy of Sciences* 94:3811-3816.
- Cheng, C.-H. C., and Chen, L. 1999. Evolution of an antifreeze glycoprotein. *Nature* 401:443-444.
- Cilia, V., Lafay, B., Christen, R. 1996. Sequence heterogeneities among 16S ribosomal RNA sequences, and their effect on phylogenetic analyses at the species level. *Molecular*

Biology and Evolution 13(3):451-461.

Clarke, L., and Carbon, J. 1976. A colony bank containing synthetic Col E1 hybrid plasmids representative of the entire *E. coli* genome. *Cell* 9:91-99.

Cohen, T., Gertler, A., and Birk, Y. 1981. Pancreatic proteolytic enzymes from carp (*Cyprinus carpio*) – I. Purification and physical properties of trypsin, chymotrypsin, elastase and carboxypeptidase B. *Comparative Biochemistry and Physiology* 69B:639-646.

Corey, D. R., and Craik, C. S. 1993. Trypsin: a model enzyme for the introduction of novel properties into proteins. In “Innovations in Proteases and Their Inhibitors” (F. X. Avilés, Ed.), pp. 425-444, Walter De Gruyter, Berlin.

Cox, T. F., and Cox, M. A. A. 1994. “Multidimensional Scaling,” Chapman & Hall, London.

Craik, C. S., Rutter, W. J. and Fletterick, R. 1983. Splice junctions: association with variation in protein structure. *Science* 220:1125-1129.

Craik, C. S., Sprang, S., Fletterick, R. and Rutter, W. J. 1982. Intron-exon splice junctions map at protein surfaces. *Nature* 299:180-182.

Das, S. K., and Sinha, N. K. 1972. Goat trypsin: purification, properties, and amino acid composition. *Indian Journal of Biochemistry and Biophysics* 9:287-292.

Davie, E. W., and Neurath, H. 1955. Identification of a peptide released during autocatalytic activation of trypsinogen. *Journal of Biological Chemistry* 212:515-529.

Davis, C. A., Riddell, D. C., Higgins, M. J., Holden, J. A., and White, B. N. 1985. A gene family in *Drosophila melanogaster* for trypsin-like enzymes. *Nucleic Acids Research* 13(18):6605-6619.

Davison, A. J. 1991. Experience in shotgun sequencing a 134 kilobase pair DNA molecule. *DNA Sequence* 1:389-394.

Dawkins, R. 1990. “The Selfish Gene,” Oxford University Press, Oxford.

Dawkins, R. 1996. “The Blind Watchmaker: Why the Evidence of Evolution Reveals a Universe Without Design,” W. W. Norton & Company, New York.

- Dayhoff, M. O. 1979. "Atlas of Protein Sequence and Structure, Volume 5, Supplement 3," National Biomedical Research Foundation, Washington, D.C.
- de Haën, C., Neurath, H., and Teller, D. C. 1975. The phylogeny of trypsin-related serine proteases and their zymogens. New methods for the investigation of distant evolutionary relationships. *Journal of Molecular Biology* 92:225-259.
- de Haën, C., Walsh, K. A., and Neurath, H. 1977. Isolation and amino-terminal sequence analysis of a new pancreatic trypsinogen of the African lungfish *Protopterus aethiopicus*. *Biochemistry* 16(20):4421-4425.
- Deininger, P. L. 1983. Random subcloning of sonicated DNA: application to shotgun DNA sequence analysis. *Analytical Biochemistry* 129:216-223.
- Demel, K., and Rutkowicz, S. 1958. "The Barents Sea," Wydawnictwo Morskie, Gdynia, Poland.
- DeRisi, J. L., Iyer, V. R., and Brown, P. O. 1997. Exploring the metabolic and genetic control of gene expression on a genomic scale. *Science* 278:680-686.
- Desnuelle, P., and Roverly, M. 1961. The proteins of the exocrine pancreas. *Advances in Protein Chemistry* 16:139-195.
- Dickerson, R. E. 1971. The structure of cytochrome c and the rates of molecular evolution. *Journal of Molecular Evolution* 1:26-45.
- Douglas, S. E., and Gallant, J. W. 1998. Isolation of cDNAs for trypsinogen from the winter flounder, *Pleuronectes americanus*. *Journal of Marine Biotechnology* 6:214-219.
- Dover, G. A., Linares, A. R., Bowen, T., and Hancock, J. M. 1993. Detection and quantification of concerted evolution and molecular drive. *Methods in Enzymology* 224:525-541.
- Dover, G., Brown, S., Coen, E., Dallas, J., Strachan, T., and Trick, M. 1982. The dynamics of genome evolution and species differentiation. In "Genome Evolution" (Dover, G. A., and Flavell, R. B., Eds.), pp. 343-372, Academic Press, London.
- Duman, J. G., and DeVries, A. L. 1974. Freezing resistance in winter flounder

Pseudopleuronectes americanus (note on *Trematomus borchgrevinki*). *Nature* 247(5438):237-238.

Eastman, J. T. 1991. The fossil and modern fish faunas of Antarctica: evolution and diversity. In "Biology of Antarctic Fish" (Prisco, G. di, Maresca, B., and Tota, B., Eds.), pp. 116-130, Springer-Verlag, New York.

Eddy, S. R., Mitchison, G., and Durbin, R. 1995. Maximum discrimination hidden Markov models of sequence consensus. *Journal of Computational Biology* 2(1):9-24.

Edelman, G. M., and Gally, J. A. 1970. Arrangement and evolution of eukaryotic genes. In "The Neurosciences: Second Study Program" (Schmitt, F. O., Ed.), pp. 962-972, Rockefeller University Press, New York.

Edwards, A., and Caskey, T. 1991. Closure strategies for random DNA sequencing. *Methods: A Companion to Methods in Enzymology*. 3(1):41-47.

Edwards, A., Voss, H., Rice, P., Civitello, A., Stegemann, J., Schwager, C., Zimmerman, J., Erfle, H., Caskey, T., and Ansorge, W. 1990. Automated DNA sequencing of the human HPRT locus. *Genomics*. 6:593-608.

Eigen, M., Winkler-Oswatitsch, R. and Dress, A. 1988. Statistical geometry in sequence space: a method of quantitative comparative sequence analysis. *Proceedings of the National Academy of Sciences* 85:5913-5917.

Emi, M., Nakamura, Y., Ogawa, M., Yamamoto, T., Nishide, T., Mori, T., and Matsubara, K. 1986. Cloning, characterization and nucleotide sequences of two cDNAs encoding human pancreatic trypsinogens. *Gene* 41(2-3):305-310.

Evans, G. A. 1991. Combinatoric strategies for genome mapping. *Bioessays* 13(1):39-44.

Everitt, B. S. 1993. "Cluster Analysis," Halsted Press, New York.

Everitt, B. S., and Dunn, G. 1991. "Applied Multivariate Data Analysis," Edward Arnold, London.

Evnin, L. B., Vásquez, J. R., and Craik, C. S. 1990. Substrate specificity of trypsin investigated by using a genetic selection. *Proceedings of the National Academy of*

Sciences 87(17):6659-6663.

Fabret, C., Quentin, Y., Chapal, N., Guiseppi, A., Haiech, J. and Denizot, F. 1996. Integrated mapping and sequencing of a 115 kb DNA fragment from *Bacillus subtilis*: sequence analysis of a 21 kb segment containing the sigL locus. *Microbiology* 142:3089-3096.

Farber, S., Shwachman, H., and Maddock, C. H. 1943. Pancreatic function and disease in early life. 1. Pancreatic enzyme activity and the coeliac syndrome. *Journal of Clinical Investigation* 22:827-838.

Feder, J. N., Gnirke, A., Thomas, W., Tsuchihashi, Z., Ruddy, D. A., Basava, A., Dormishian, F., Domingo, R., Jr., Ellis, M. C., Fullan, A., Hinton, L. M., Jones, N. L., Kimmel, B. E., Kronmal, G. S., Lauer, P., Lee, V. K., Loeb, D. B., Mapa, F. A., McClelland, E., Meyer, N. C., Mintier, G. A., Moeller, N., Moore, T., Morikang, E., Prass, C. E., Quintana, L., Starnes, S. M., Schatzman, R. C., Brunke, K. J., Drayna, D. T., Risch, N. J., and Wolff, R. K. 1996. A novel MHC class I-like gene is mutated in patients with hereditary haemochromatosis. *Nature Genetics* 13:399-408.

Feller, G., and Gerday, C. 1997. Psychrophilic enzymes: molecular basis of cold adaptation. *Cellular and Molecular Life Sciences* 53(10):830-841.

Felsenstein, J. 1983. Inferring evolutionary trees from DNA sequences. In "Statistical Analysis of DNA Sequence Data" (Weir, B. S., Ed.), pp. 133-150, Marcel Dekker, New York.

Felsenstein, J. 1993. PHYLIP (Phylogeny inference package) version 3.5c. Distributed by the author, Department of Genetics, University of Washington, Seattle.

Felsenstein, J., and Churchill, G. A. 1996. A Hidden Markov Model approach to variation among sites in rate of evolution. *Molecular Biology and Evolution* 13(1):93-104.

Fields, C., Adams, M. D., White, O. and Venter, J. C. 1994. How many genes in the human genome? *Nature Genetics* 7:345-346.

Fisher, R. A. 1940. On the similarity of the distributions found for the test of harmonic significance in harmonic analysis, and in Steven's problem in geometrical probability. *Annals of Eugenics* 10:14-17.

- Fitch, W. M. 1970. Distinguishing homologous from analogous proteins. *Systematic Zoology* 19:99-113.
- Flatto, L., and A. G. Konheim. 1962. The Random division of an interval and the random covering of a circle. *SIAM Review* 43:211-222.
- Fleischmann, R. D., Adams, M. D., White, O., Clayton, R. A., Kirkness, E. F., Kerlavage, A. R., Bult, C. J., Tomb, J. F., Dougherty, B. A., Merrick, J. M., McKenney, K., Sutton, G., FitzHugh, W., Fields, C., Gocayne, J. D., Scott, J., Shirley, R., Liu, L.-I., Glodek, A., Kelley, J. M., Weidman, J. F., Phillips, C. A., Spriggs, T., Hedblom, E., Cotton, M. D., Utterback, T. R., Hanna, M. C., Nguyen, D. T., Saudek, D. M., Brandon, R. C., Fine, L. D., Fritchman, J. L., Fuhrman, J. L., Geoghagen, N. S. M., Gnehm, C. L., McDonald, L. A., Small, K. V., Fraser, C. M., Smith, H. O. and Venter, J. C. 1995. Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. *Science* 269:496-512.
- Fletcher, T. S., Alhadeff, M., Craik, C. S., and Langman, C. 1987. Isolation and characterization of a cDNA encoding rat cationic trypsinogen. *Biochemistry* 26:3081-3086.
- Fraser, C. M., Gocayne, J. D., White, O., Adams, M. D., Clayton, R. A., Fleischmann, R. D., Bult, C. J., Kerlavage, A. R., Sutton, G., Kelley, J. M., Fritchman, J. L., Weidman, J. F., Small, K. V., Sandusky, M., Fuhrmann, J., Nguyen, D., Utterback, T. R., Saudek, D. M., Phillips, C. A., Merrick, J. M., Tomb, J.-F., Dougherty, B. A., Bott, K. F., Hu, P.-C., Lucier, T. S., Peterson, S. N., Smith, H. O., Hutchison, C. A. and Venter, J. C. 1995. The minimal gene complement of *Mycoplasma genitalium*. *Science* 270:397-403.
- Galper, A. R., and Brutlag, D. L. 1993. Computational simulations of biological systems. In "Biocomputing: Informatics and Genome Projects" (Smith, D. W., Ed.), pp. 269-305, Academic Press, San Diego.
- Genicot, S., Feller, G., and Gerday, C. 1988. Trypsin from Antarctic fish (*Paranotothenia magellanica* Forster) as compared with trout (*Salmo gairdneri*) trypsin. *Comparative Biochemistry and Physiology* 90B(3):601-609.
- Genicot, S., Rentier-Delrue, F., Edwards, D., vanBeeumen, J., and Gerday, C. 1996. Trypsin and trypsinogen from an Antarctic fish: molecular basis of cold adaptation. *Biochimica et*

Biophysica Acta 1298(1):45-57.

- Gish, W., and States, D. J. 1993. Identification of protein coding regions by database similarity search. *Nature Genetics* 3:266-272.
- Gon, O., and Heemstra, P. C. 1990. "Fishes of the Southern Ocean," J. L. B. Smith Institute of Ichthyology, Grahamstown, South Africa.
- Gonick, L., and Wheelis, M. 1991. "The Cartoon Guide to Genetics," Harper Perennial, New York.
- Green, E. D., and Green, P. 1991. Sequence tagged site (STS) content mapping of human chromosomes: theoretical considerations and early experiences. *PCR Methods and Applications* 1:77-90.
- Green, N. M., and Neurath, H. 1954. Proteolytic enzymes. In "The Proteins, Volume II" (Neurath, H., and Bailey, K., Eds.), pp. 1057-1198, Academic Press, New York.
- Gudmundsdóttir, Á., Gudmundsdóttir, E., Óskarsson, S., Bjarnason, J. B., Eakin, A. K., and Craik, C. S. 1993. Isolation and characterization of cDNAs from Atlantic cod encoding two different forms of trypsinogen. *European Journal of Biochemistry* 217:1091-1097.
- Gusfield, D. 1997. "Algorithms on Strings, Trees, and Sequences: Computer Science and Computational Biology," Cambridge University Press, Cambridge.
- Haars, R., Kronenberg, M., Gallatin, W. M., Weissman, I. L., Owen, F. L. and Hood, L. 1986. Rearrangement and expression of T cell antigen receptor and gamma genes during thymic development. *Journal of Experimental Medicine* 164:1-24.
- Hagmann, M. 1997. RAGged repair: what's new in V(D)J recombination. *Biological Chemistry* 378:815-819.
- Hall, P. 1988. "Introduction to the Theory of Coverage Processes," John Wiley & Sons, New York.
- Hartley, B. S. 1970. Homologies in serine proteinases. *Philosophical Transactions of the Royal Society of London. Series B. Sci* 257:77-87.

- Hartley, B. S. 1979. Evolution of enzyme structure. *Proceedings of the Royal Society of London. Series B* 205:443-452.
- Haskins, C. H. 1923. "The Rise of Universities," Cornell University Press, London.
- Hedstrom, L., Farr-Jones, S., Kettner, C. A., and Rutter, W. J. 1994a. Converting trypsin to chymotrypsin: ground-state binding does not determine substrate specificity. *Biochemistry* 33:8764-8769.
- Hedstrom, L., Perona, J. J., and Rutter, W. J. 1994b. Converting trypsin to chymotrypsin: residue 172 is a substrate specificity determinant. *Biochemistry* 33:8757-8763.
- Hedstrom, L., Szilagyi, L., and Rutter, W. J. 1992. Converting trypsin to chymotrypsin: the role of surface loops. *Science* 255:1249-1253.
- Heiter, P. and Boguski, M. 1997. Functional genomics: it's all how you read it. *Science* 278:601-602.
- Henikoff, S., Greene, E. A., Pietrokovski, S., Bork, P., Attwood, T. K. and Hood, L. 1997. Gene families: the taxonomy of protein paralogs and chimeras. *Science* 278:609-614.
- Hermanson, M. A., Ericsson, L. H., Neurath, H., and Walsh, K. A. 1973. Determination of the amino acid sequence of porcine trypsin by sequenator analysis. *Biochemistry* 12(17):3146-3153.
- Hewett-Emmett, D., Czelusniak, J. and Goodman, M. 1981. The evolutionary relationship of the enzymes involved in blood coagulation and hemostasis. *Annals of the New York Academy of Sciences* 370:511-527.
- Higashimoto, Y. and Liddle, R. A. 1993. Isolation and characterization of the gene encoding rat glucose-dependent insulinotropic peptide. *Biochemical and Biophysical Research Communications* 193:182-190.
- Hillis, D. M. 1996. Inferring complex phylogenies. *Nature* 383:130-131.
- Himmelreich, R., Hilbert, H., Plagens, H., Pirkl, E., Li, B. C., and Herrmann, R. 1996. Complete sequence analysis of the genome of the bacterium *Mycoplasma pneumoniae*. *Nucleic Acids Research* 24:4420-4449.

- Hjelmeland, K., and Raa, J. 1982. Characteristics of two trypsin type isozymes isolated from the Arctic fish capelin (*Mallotus villosus*). *Comparative Biochemistry and Physiology* 71B:557-562.
- Hofer, R., Schiestl, W., Gattringer, A., Ladurner, H., and Wieser, W. 1975. Trypsin isozymes of some ectothermic vertebrates. *Journal of Comparative Physiology* 101:111-119.
- Hofmann, T. 1964. The purification and properties of fragments of trypsinogen obtained by cyanogen bromide cleavage. *Biochemistry* 3:356-364.
- Holland, P. W. and Garcia-Fernàndez, J. 1996. Hox genes and chordate evolution. *Developmental Biology* 173:382-395.
- Hollingshead, S. K., Arnold, J., Readdy, T. L., Bessen, D. E. 1994. Molecular evolution of a multigene family in group A streptococci. *Molecular Biology and Evolution* 11(2):208-219.
- Hood, L., and Hunkapiller, T. 1991. Molecular evolution and the immunoglobulin gene superfamily. In "Evolution of Life" (Osawa, S., and Honjo, T., Eds.), pp. 123-243, Springer-Verlag, Tokyo.
- Hood, L., Campbell, J. H., Elgin, S. C. 1975. The organization, expression, and evolution of antibody genes and other multigene families. *Annual Review of Genetics* 9:305-353.
- Huber, R. and Bode, W. 1978. Structural basis of the activation and action of trypsin. *Accounts of Chemical Research* 11:114-122.
- Hunkapiller, T. and Hood, L. 1989. Diversity of the immunoglobulin gene superfamily. *Advances in Immunology* 44:1-63.
- Hunkapiller, T., Gorman, J., Koop, B. F. and Hood, L. 1989. Implications of the diversity of the immunoglobulin gene superfamily. *Cold Spring Harbor Symposia on Quantitative Biology* 54(1):15-29.
- Iwabe, N., Kuma, K., and Miyata, T. Evolution of gene families and relationship with organismal evolution: rapid divergence of tissue-specific genes in the early evolution of chordates. *Molecular Biology and Evolution* 13(3):483-493.

- Jin, L., and Nei, M. 1990. Limitations of the evolutionary parsimony method of phylogenetic analysis. *Molecular Biology and Evolution* 7(1):82-102.
- Jones, D. T., Taylor, W. R. and Thornton, J. M. 1994. A mutation data matrix for transmembrane proteins. *FEBS Letters* 339:269-275.
- Jones, D. T., Taylor, W. R., and Thornton, J. M. 1992. The rapid generation of mutation data matrices from protein sequences. *Computer Applications in the Biosciences* 8(3):275-282.
- Jukes, T. H., and Cantor, C. R. 1969. Evolution of protein molecules. In "Mammalian Protein Metabolism" (H. N. Munro, ed.), pp. 21-132, Academic Press, New York.
- Jurasek, L., Fackre, D. and Smillie, L. B. 1969. Remarkable homology about the disulfide bridges of a trypsin-like enzyme from *Streptomyces griseus*. *Biochemical and Biophysical Research Communications* 37:99-105.
- Jurášek, L. and Smillie, L. B. 1973. The amino acid sequence of *Streptomyces griseus* trypsin. I. The peptic peptides. *Canadian Journal of Biochemistry* 51:1077-1088.
- Kaláč, J. 1978. Studies on herring (*Clupea harengus* L.) and capelin (*Mallotus villosus* L.) pyloric caeca proteases II: characterization of the anionic fraction of trypsins. *Biológia (Bratislava)* 33(9):701-710.
- Kaneko, T., Sato, S., Kotani, H., Tanaka, A., Asamizu, E., Nakamura, Y., Miyajima, N., Hirosawa, M., Sugiura, M., Sasamoto, S., Kimura, T., Hosouchi, T., Matsuno, A., Muraki, A., Nakazaki, N., Naruo, K., Okumura, S., Shimpo, S., Takeuchi, C., Wada, T., Watanabe, A., Yamada, M., Yasuda, M. and Tabata, S. 1996. Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis* sp. strain PCC6803. II. Sequence determination of the entire genome and assignment of potential protein-coding regions. *DNA Research* 3:109-136.
- Kang, J., Wiegand, U., and Müller-Hill, B. 1992. Identification of cDNAs encoding two novel rat pancreatic serine proteases. *Gene* 110:181-187.
- Karlin, K., Blaisdell, B. E., and Bucher, P. 1992. Quantile distributions of amino acid usage in protein classes. *Protein Engineering* 5(8):729-738.

- Kasai, H., Isono, S., Mineno, J., Akiyama, H., Kurnit, D.M., Berg, D.E., and Isono, K. 1992. Efficient large-scale sequencing of the *Escherichia coli* genome: implementation of a transposon-and PCR-based strategy for the analysis of ordered phage clones. *Nucleic Acids Research* 20(24):6509-6515.
- Kauffman, D. L. 1965. The disulphide bridges of trypsin. *Journal of Molecular Biology* 12:929-932.
- Kim, J. C., Cha, S. H., Jeong, S. T., Oh, S. K. and Byun, S. M. 1991. Molecular cloning and nucleotide sequence of *Streptomyces griseus* trypsin gene. *Biochemical and Biophysical Research Communications* 181:707-713.
- Kimura, M., and Ota, T. 1971. On the rate of molecular evolution. *Journal of Molecular Evolution* 1:1-17.
- Koide, A., and Matsuoka, Y. 1971. Studies on fin whale pancreatic proteases - IV. Properties of purified fin whale cationic trypsin. *Bulletin of the Japanese Society of Scientific Fisheries* 37(6):545-552.
- Koonin, S. E. 1998. An independent perspective on the human genome project. *Science* 279:36-37.
- Koop, B. F., Rowen, L., Chen, W.-Q., Deshpande, P., Lee, H., and Hood, L. 1993. Sequence length and error analysis of Sequenase® and automated *Taq* cycle sequencing methods. *BioTechniques* 14(3):442-447.
- Kossiakoff, A. A., Chambers, J. L., Kay, L. M., Stroud, R. M. 1977. Structure of bovine trypsinogen at 1.9 Å resolution. *Biochemistry* 16(4):654-664.
- Lander, E. S., and M. S. Waterman. 1988. Genomic mapping by fingerprinting random clones: A mathematical analysis. *Genomics* 2:231-239.
- Lange, K., and Boehnke, M. 1982. How many polymorphic genes will it take to span the human genome? *American Journal of Human Genetics* 34(6):842-845.
- Lawrence, C. B., Honda, S., Parrott, N. W., Flood, T. C., Gu, L., Zhang, L., Jain, M., Larson, S. and Myers, E. W. 1994. The genome reconstruction manager: a software environment

for supporting high-throughput DNA sequencing. *Genomics* 23:192-201.

Le Huerou, I., Wicker, C., Guilloteau, P., Toullec, R., Puigserver, A. 1990. Isolation and nucleotide sequence of cDNA clone for bovine pancreatic anionic trypsinogen. *European Journal of Biochemistry* 193:767-773.

Lévy, P. 1939. Sur la division d'un segment par des points choisis au hasard. *Comptes Rendus* 208:147-149.

Li, C., and Tucker, P. W. 1993. Exoquence DNA sequencing. *Nucleic Acids Research* 21(5):1239-1244.

Li, W.-H. 1997. "Molecular Evolution," Sinauer Associates, Sunderland, Massachusetts.

Li, W.-H., and Graur, D. 1991. "Fundamentals of Molecular Evolution," Sinauer Associates, Sunderland, Massachusetts.

Liu, G. L. 1968. "Introduction to Combinatorial Mathematics," McGraw Hill, New York.

Logsdon, J. M., Jr. and Doolittle, W. F. 1997. Origin of antifreeze protein genes: a cool tale in molecular evolution. *Proceedings of the National Academy of Sciences* 94:3485-3487.

Louvard, M. N., and Puigserver, A. 1974. On bovine and porcine anionic trypsinogens. *Biochimica and Biophysica Acta* 371:177-185.

Low, P. S., Bada, J. L., and Somero, G. N. 1973. Temperature adaptation of enzymes: roles of the free energy, the enthalpy, and the entropy of activation. *Proceedings of the National Academy of Sciences* 70(2):430-432.

Lütcke, H., Rausch, U., Vasiloudes, P., Scheele, G. A., and Kern, H. F. 1989. A fourth trypsinogen (P23) in the rat pancreas induced by CCK. *Nucleic Acids Research* 17(16):6736.

MacDonald, R. J., Stary, S. J., Swift, G. H. 1982. Two similar but nonallelic rat pancreatic trypsinogens. Nucleotide sequences of the cloned cDNAs. *Journal of Biological Chemistry* 257(16):9724-9732.

Macouzet, M., Simpson, B. K., and Lee, B. H. 1999. Cloning of fish enzymes and other fish

- protein genes. *Critical Reviews in Biotechnology* 19(3):179-196.
- Male, R., Lorens, J. B., Smalås, A. O., and Torrissen, K. R. 1995. Molecular cloning and characterization of anionic and cationic variants of trypsin from Atlantic salmon. *European Journal of Biochemistry* 232:677-685.
- Mallory, P. A., and Travis, J. 1973. Human pancreatic enzymes. Characterization of anionic human trypsin. *Biochemistry* 12(15):2847-2851.
- Maroux, S., Baratti, J. and Desnuelle, P. 1971. Purification and specificity of porcine enterokinase. *Journal of Biological Chemistry* 246:5031-5039.
- Martin-Gallardo, A., J. Lamardin, and A. Carrano. 1994. Shotgun sequencing. In "Automated DNA Sequencing and Analysis" (Adams, M., Fields, C., and Venter, J., Eds.), pp. 37-41, Academic Press, New York.
- Mayo, K. E., Cerelli, G. M., Rosenfeld, M. G. and Evans, R. M. 1985. Characterization of cDNA and genomic clones encoding the precursor to rat hypothalamic growth hormone-releasing factor. *Nature* 314:464-467.
- McKusick, V. 1997. Genomics: structural and functional studies of genomes. *Genomics* 45:244-249.
- Mikeš, O., Holeyšovský, V., Tomášek, V., and Šorm, F. 1966. Covalent structure of bovine trypsinogen. The position of the remaining amides. *Biochemical and Biophysical Research Communications* 24(3):346-352.
- The MITRE Corporation. 1997. "JASON Report JSR-97-315," The MITRE Corporation, McLean, Virginia.
- Morgan, P. H., Robinson, N. C., Walsh, K. A. and Neurath, H. 1972. Inactivation of bovine trypsinogen and chymotrypsinogen by diisopropylphosphorofluoridate. *Proceedings of the National Academy of Sciences* 69:3312-3316.
- Mugasimangalam, R. C., Zevin-Sonkin, D., Shwartzburd, J., Rozovskaya, T. A., Sobolev, I. A., Chertkov, O., Ramanathan, V., Lvovsky, L., and Ulanovsky, L. E. 1997. DNA sequencing using differential extension with nucleotide subsets (DENS). *Nucleic Acids*

Research 25(4):800-805.

Müller, W. E. G., Pancer, Z., and Rinkevich, B. 1994. Molecular cloning and localization of a novel serine protease from the colonial tunicate *Botryllus schlosseri*. *Molecular Marine Biology and Biotechnology* 3(2):70-77.

Murakami, K., and Noda, M. 1981. Studies on proteinases from the digestive organs of sardine. I. Purification and characterization of three alkaline proteinases from the pyloric caeca. *Biochimica et Biophysica Acta* 658:17-26.

Murray-Rust, J., McDonald, N. Q., Blundell, T. L., Hosang, M., Oefner, C., Winkler, F., and Bradshaw, R. A. 1993. Topological similarities in TGF-beta 2, PDGF-BB and NGF define a superfamily of polypeptide growth factors. *Structure* 1:153-159.

Nakai, K. and Kanehisa, M. 1992. A knowledge base for predicting protein localization sites in eukaryotic cells. *Genomics* 14:897-911.

Neurath, H., Dixon, G. H. 1957. Structure and activation of trypsinogen and chymotrypsinogen. *Federation Proceedings* 16:791-801.

Newberg, L. A. 1993. "Finding, Evaluating, and Counting DNA Physical Maps," Doctoral Thesis, University of California at Berkeley.

Newberg, L. A. 1996. The number of clone orderings. *Discrete Applied Mathematics* 69(3):233-245.

Northrup, J. H., Kunitz, M., and Herriott, R. 1948. "Crystalline Enzymes," Columbia University Press, New York.

Ohno, S. 1978. The significance of gene duplication in immunoglobulin evolution (epimethean natural selection and prometean evolution). In "Immunoglobulins" (Litman, G. W., and Good, R. A., Eds.), pp. 197-204, Plenum, New York.

Olson, M. V., and Green, P. 1993. Criterion for the completeness of large-scale physical maps of DNA. *Cold Spring Harbor Symposia on Quantitative Biology* 53:349-355.

Ota, T., and Kimura, M. 1971. On the constancy of the evolutionary rate of cistrons. *Journal of Molecular Evolution* 1:18-25.

- Pancer, Z., Leuck, J., Rinkevich, B., Steffen, R., Müller, I., and Müller, W. E. G. 1996. Molecular cloning and sequencing analysis of two cDNAs coding for putative anionic trypsinogens from the colonial urochordate *Botryllus schlosseri* (Ascidiacea). *Molecular Marine Biology and Biotechnology* 5(4):326-333.
- Pease, L. R., Horton, R. M., Pullen, J. K. and Yun, T. J. 1993. Unusual mutation clusters provide insight into class I gene conversion mechanisms. *Molecular and Cellular Biology* 13:4374-4381.
- Perona, J. J., and Craik, C. S. 1995. Structural basis of substrate specificity in the serine proteases. *Protein Science* 4(3):337-360.
- Perona, J. J., and Craik, C. S. 1997. Evolutionary divergence of substrate specificity within the chymotrypsin-like serine protease fold. *Journal of Biological Chemistry* 272(48):29987-29990.
- Perona, J. J., Evin, L. B., and Craik, C. S. 1993. A genetic selection elucidates structural determinants of arginine versus lysine specificity in trypsin. *Gene* 137(1):121-126.
- Petes, T. D. and Hill, C. W. 1988. Recombination between repeated genes in microorganisms. *Annual Review of Genetics* 22:147-168.
- Petes, T. D., Malone, R. E., and Symington, L. S. 1991. Recombination in yeast. In "The Molecular and Cellular Biology of the Yeast *Saccharomyces*: Genome Dynamics, Protein Synthesis, and Energetics, Volume I" (Broach, J., Jones, E., and Pringle, J., Eds.), pp. 407-521, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.
- Pinsky, S. D., LaForge, K. S., and Scheele, G. 1985. Differential regulation of trypsinogen mRNA translation: full-length mRNA sequences encoding two oppositely charged trypsinogen isoenzymes in the dog pancreas. *Molecular and Cellular Biology* 5(10):2669-2676.
- Port, E., Sun, F., Martin, D., and Waterman, M. S. 1995. Genomic mapping by end-characterized random clones: a mathematical analysis. *Genomics* 26:84-100.
- Rapoport, T. A. 1990. Protein transport across the ER membrane. *Trends in Biochemical Sciences* 15:355-358.

- Raport, C. J., Schweickart, V. L., Chantry, D., Eddy, R. L., Jr., Shows, T. B., Godiska, R. and Gray, P. W. 1996. New members of the chemokine receptor gene family. *Journal of Leukocyte Biology* 59:18-23.
- Ravitch, M. M. 1973. Ectopic Pancreas – not so rare, not so innocent. *Medical Times* 101(6):57-59.
- Read, R. J., and James, M. N. 1988. Refined crystal structure of *Streptomyces griseus* trypsin at 1.7 Å resolution. *Journal of Molecular Biology* 200:523-551.
- Reeck, G. R., and Neurath, H. 1972. Pancreatic trypsin from the African lungfish. *Biochemistry* 11(4):503-510.
- Reynaud, C. A., Anquez, V., Dahan, A., and Weill, J. C. 1985. A single rearrangement event generates most of the chicken immunoglobulin light chain diversity. *Cell* 40:283-291.
- Richards, S., Muzny, D. M., Civitello, D. M., Lu, F., and Gibbs, R. A. 1994. Sequence map gaps and directed reverse sequencing for the completion of large sequencing projects. In “Automated DNA Sequencing and Analysis” (Adams, M., Fields, C., and Venter, J., Eds.), pp. 191-198, Academic Press, New York.
- Roach, J. C. 1995. Random subcloning. *Genome Research* 5:464-473.
- Roach, J. C., Boysen, C., Wang, K., and Hood, L. 1995. Pairwise end sequencing: a unified approach to genomic mapping and sequencing. *Genomics* 26:345-353.
- Roach, J. C., Wang, K., Gan, L., Hood, L. 1997. The molecular evolution of the vertebrate trypsinogens. *Journal of Molecular Evolution* 45(6):640-652.
- Robbins, H. E. 1944. On the measure of a random set. *Annals of Mathematical Statistics* 15:70-74.
- Robbins, H. E. 1945. On the measure of a random Set. II. *Annals of Mathematical Statistics* 16:342-347.
- Rogers. 1985. Exon shuffling and intron insertion in serine protease genes. *Nature* 315:458-459.

- Rowen, L., and Koop, B. F. 1994. Zen and the art of large-scale genomic sequencing. In "Automated DNA Sequencing and Analysis" (Adams, M., Fields, C., and Venter, J., Eds.), pp. 167-174, Academic Press, New York.
- Rowen, L., Koop, B. F., and Hood, L. 1996. The complete 685-kilobase DNA sequence of the human T cell receptor locus. *Science* 272:1755-1762.
- Rowen, L., Mahairas, G., and Hood, L. 1997. Sequencing the human genome. *Science* 278:605-607.
- Ryan, C. A., Clary, J. J., and Tomimatsu, Y. 1965. Chicken chymotrypsin and turkey trypsin. *Archives of Biochemistry and Biophysics* 110:175-183.
- Rypniewski, W. R., Perrakis, A., Vorgias, C. E., and Wilson, K. S. 1994. Evolutionary divergence and conservation of trypsin. *Protein Engineering* 7(1):57-64.
- Scheele, G., Bartelt, D. and Bieger, W. 1981. Characterization of human exocrine pancreatic proteins by two-dimensional isoelectric focusing/sodium dodecyl sulfate gel electrophoresis. *Gastroenterology* 80:461-473.
- Schick, J., Kern, H. and Scheele, G. 1984. Hormonal stimulation in the exocrine pancreas results in coordinate and anticoordinate regulation of protein synthesis. *Journal of Cell Biology* 99:1569-1574.
- Schneider, T. D. and Stephens, R. M. 1990. Sequence logos: a new way to display consensus sequences. *Nucleic Acids Research* 18:6097-6100.
- Schyns, R., Bricteux-Grégoire, S., and Florkin, M. 1969. Purification, propriétés et séquence N-terminale du trypsinogène de mouton. *Biochimica et Biophysica Acta* 175:97-112.
- Seboun, E., Houghton, L., Hatem, C. J., Jr., Lincoln, R. and Hauser, S. L. 1993. Unusual organization of the human T-cell receptor beta-chain gene complex is linked to recombination hotspots. *Proceedings of the National Academy of Sciences* 90:5026-5029.
- Seboun, E., Joshi, N. and Hauser, S. L. 1992. Haplotypic origin of beta-chain genes expressed by human T-cell clones. *Immunogenetics* 36:363-368.

- Seboun, E., Robinson, M. A., Kindt, T. J. and Hauser, S. L. 1989. Insertion/deletion-related polymorphisms in the human T cell receptor beta gene complex. *Journal of Experimental Medicine* 170:1263-1270.
- Shi, Y.-B., and Brown, D. D. 1990. Developmental and thyroid hormone-dependent regulation of pancreatic genes in *Xenopus laevis*. *Genes and Development* 4(7):1107-1113.
- Siegel, A. F. 1979. Asymptotic coverage distributions on the circle. *Annals of Probability* 74:651-661.
- Siegel, A. F., and L. Holst. 1982. Covering the circle with random arcs of random sizes. *Journal of Applied Probability* 19:373-381.
- Siegel, A. F., Roach, J. C., and van den Engh, G. 1998a. Expectation and variance of true and false fragment matches in DNA restriction mapping. *Journal of Computational Biology* 5(1):101-111.
- Siegel, A. F., Roach, J. C., Magness, C., Thayer, E., and van den Engh, G. 1998b. Optimization of restriction fragment DNA mapping. *Journal of Computational Biology* 5(1):113-126.
- Siegel, A. F., Trask, B., Roach, J. C., Mahairas, G. G., Hood, L., and van den Engh, G. 1999. Analysis of sequence-tagged-connector strategies for DNA sequencing. *Genome Research* 9(3):297-307.
- Siemieniak, D. L., Sieu, L. C., and Slightom, J. L. 1991. Strategy and methods for directly sequencing cosmid clones. *Analytical Biochemistry* 192:441-448.
- Simpson, B. K., and Haard, N. F. 1984a. Purification and characterization of trypsin from the Greenland cod (*Gadus ogac*). I. Kinetic and thermodynamic characteristics. *Canadian Journal of Biochemistry and Cell Biology* 62:894-900.
- Simpson, B. K., and Haard, N. F. 1984b. Trypsin from Greenland cod, *Gadus ogac*. Isolation and comparative properties. *Comparative Biochemistry and Physiology* 79B(4):613-622.
- Smalås, A. O., Heimstad, E. S., Hordvik, A., Willassen, N. P. and Male, R. 1994. Cold adaption of enzymes: structural comparison between salmon and bovine tryptins. *Proteins*

20:149-166.

- Smith, M. W., Holmsen, A. L., Wei, Y. H., Peterson, M., and Evans, G. A. 1994. Genomic sequence sampling: A strategy for high resolution sequence-based physical mapping of complex genomes. *Nature Genetics* 7(1):40-47.
- Smyth, M. J., O'Connor, M. D. and Trapani, J. A. 1996. Granzymes: a variety of serine protease specificities encoded by genetically distinct subfamilies. *Journal of Leukocyte Biology* 60:555-562.
- Solomon, H. 1978. "Geometric Probability," Society for Industrial and Applied Mathematics, Philadelphia, Pennsylvania.
- Staden, R. 1980. A new computer method for the storage and manipulation of DNA gel reading data. *Nucleic Acids Research* 8:3673-3694.
- Steiner, J. M., Medinger, T. L., and Williams, D. A. 1997. Purification and partial characterization of feline trypsin. *Comparative Biochemistry and Physiology* 116B(1):87-93.
- Steinhilber, W., Poensgen, J., Rausch, U., Kern, H. F. and Scheele, G. A. 1988. Translational control of anionic trypsinogen and amylase synthesis in rat pancreas in response to caerulein stimulation. *Proceedings of the National Academy of Sciences* 85:6597-6601.
- Stevens, W. L. 1939. Solution to a geometrical problem in probability. *Annals of Eugenics* 9:315-320.
- Stevenson, B. J., Hagenbüchle, O., and Wellauer, P. K. 1986. Sequence organisation and transcriptional regulation of the mouse elastase II and trypsin genes. *Nucleic Acids Research* 14(21):8307-8330.
- Stewart, C.-B., and Wilson, A. C. 1987. Sequence convergence and functional adaptation of stomach lysozymes from foregut fermenters. *Cold Spring Harbor Symposia on Quantitative Biology* 52:891-899.
- Stoltzfus, A., Logsdon, J. M., Jr., Palmer, J. D. and Doolittle, W. F. 1997. Intron "sliding" and the diversity of intron positions. *Proceedings of the National Academy of Sciences* 94:10739-10744.

- Stubbs, L. 1992. Long-range walking techniques in positional cloning strategies. *Mammalian Genome* 3:127-142.
- Suyama, M., Matsuo, Y. and Nishikawa, K. 1997. Comparison of protein structures using 3D profile alignment. *Journal of Molecular Evolution* 44(Suppl. 1):S163-173.
- Sweet, R. M., Wright, H. T., Janin, J., Chothia, C. H., and Blow, D. M. 1974. Crystal structure of the complex of porcine trypsin with soybean trypsin inhibitor (Kunitz) at 2.6-Å resolution. *Biochemistry* 13(20):4212-4228.
- Szabo, C. I., Fransisco, L. V., Roach, J., Argonza, R., Wagner, L. A., Ostrander, E. A., King, M.-C. 1996. Human, canine, and murine BRCA1 genes: sequence comparison among species. *Human Molecular Genetics* 5(9):1289-1298.
- Tani, T., Kawashima, I., Mita, K., Takiguchi, Y. 1990. Nucleotide sequence of the human pancreatic trypsinogen III cDNA. *Nucleic Acids Research* 18(6):1631.
- Tatusov, R. L., Koonin, E. V. and Lipman, D. J. 1997. A genomic perspective on protein families. *Science* 278:631-637.
- Taylor, W. R., and Jones, D. T. 1993. Deriving an amino acid distance matrix. *Journal of Theoretical Biology* 164(1):65-83.
- Thoeni, R. F., and Gedgaudas, R. K. 1980. Ectopic pancreas: usual and unusual features. *Gastrointestinal Radiology* 5(1):37-42.
- Thompson, J. D., Higgins, D. G. and Gibson, T. J. 1994. *CLUSTAL W*: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. *Nucleic Acids Research* 22:4673-4680.
- Thorne, J. L., Goldman, N., and Jones, D. T. 1996. Combining protein evolution and secondary structure. *Molecular Biology and Evolution* 13(5):666-673.
- Titani, K., Ericsson, L. H., Neurath, H., and Walsh, K. A. 1975. Amino acid sequence of dogfish trypsinogen. *Biochemistry* 14(7):1358-1366.
- Titani, K., Sasagawa, T., Woodbury, R. G., Ericsson, L. H., Dörsam, H., Kraemer, M.,

- Neurath, H., and Zwilling, R. 1983. Amino acid sequence of crayfish (*Astacus fluviatilis*) trypsin. *Biochemistry* 22:1459-1465.
- Townes, P. L. 1972. Trypsinogen deficiency and other proteolytic deficiency diseases. *Birth Defects: Original Article Series* 8(2):95-101.
- Travis, J. 1968. Studies on the active site of sheep trypsin. *Biochemical and Biophysical Research Communications* 30(6):730-734.
- Travis, J., and Liener, I. E. 1965. The crystallization and partial characterization of porcine trypsin. *The Journal of Biological Chemistry* 240(5):1962-1966.
- Travis, J., and Roberts, R. C. 1969. Human trypsin. Isolation and physical-chemical characterization. *Biochemistry* 8(7):2884-2889.
- Uchida, N. 1973. Doctoral Thesis, Hokkaido University.
- Vandermeers, A., and Cristophe, J. 1969. Purification, activite enzymatique potentielle et composition en acides amines du chymotrypsinogene et des trypsinogenes du pancreas du rat. *Biochimica et Biophysica Acta* 188:101-112.
- Venter, J. C., Smith, H. O., and Hood, L. 1996. A new strategy for genome sequencing. *Nature* 381(6581):364-366.
- Vingron, M., and von Haeseler, A. 1997. Towards integration of multiple alignment and phylogenetic tree construction. *Journal of Computational Biology* 4(1):23-34.
- Viswanatha, T., Lawson, W. B., and Witkop, B. 1960. The action of N-bromosuccinimide on trypsinogen and its derivatives. *Biochimica et Biophysica Acta* 40:216-224.
- von Heijne, G. 1985. Signal sequences. The limits of variation. *Journal of Molecular Biology* 184:99-105.
- Voytek, P., and Gjessing, E. C. 1971. Studies of an anionic trypsinogen and its active enzyme from porcine pancreas. *The Journal of Biological Chemistry* 246(2):508-516.
- Wachtel, S. S., and Tiersch, T. R. 1993. Variations in genome mass. *Comparative Biochemistry and Physiology* 104B(2):207-213.

- Walsh, K. A. 1970. Trypsinogens and trypsins of various species. *Methods in Enzymology* 19:41-63.
- Walsh, K. A., and Neurath, H. 1964. Trypsinogen and chymotrypsinogen as homologous proteins. *Proceedings of the National Academy of Sciences* 52:884-889.
- Walsh, K. A., and Wilcox, P. E. 1970. Serine proteases. *Methods in Enzymology* 19:31-41.
- Wang, K., Gan, L., Lee, I., and Hood, L. 1995. Isolation and characterization of the chicken trypsinogen gene family. *Biochemical Journal* 307:471-479.
- Waterman, M. S. 1995. "Introduction to Computational Biology: Maps, sequences, and genomes," Chapman & Hall, London.
- Whitcomb, D. C., Gorry, M. C., Preston, R. A., Furrey, W., Sossenheimer, M. J., Ulrich, C. D., Martin, S. P., Gates, L. K., Amann, S. T., Toskes, P. P., Liddle, R., McGrath, K., Uomo, G., Post, J. C., and Ehrlich, G. D. 1996. Hereditary pancreatitis is caused by a mutation in the cationic trypsinogen gene. *Nature Genetics* 14:141-145.
- Whitworth, W. A. 1897a. "Choice and Chance," Cambridge University Press, Cambridge.
- Whitworth, W. A. 1897b. "Exercises in Choice and Chance," Cambridge University Press, Cambridge.
- Wiegand, U., Corbach, S., Minn, A., Kang, J., and Müller-Hill, B. 1993. Cloning of the cDNA encoding human brain trypsinogen and characterization of its product. *Gene* 136:167-175.
- Wiemann, S., Stegemann, J., Zimmermann, J., Voss, H., Benes, V., and Ansorge, W. 1996. "Doublex" fluorescent DNA sequencing: two independent sequences obtained simultaneously in one reaction with internal labeling and unlabeled primers. *Analytical Biochemistry* 234(2):166-174.
- Wills, C. 1991. "Exons, Introns, and Talking Genes: The Science Behind the Human Genome Project," pp. 92-95, Harper Collins, New York.
- Wu, C. F. J. 1986. Jackknife, bootstrap and other resampling plans in regression analysis. *Annals of Statistics* 14:1261-1295.

- Yang, Z. 1993. Maximum-likelihood estimation of phylogeny from DNA sequences when substitution rates differ over sites. *Molecular Biology and Evolution* 10(6):1396-1401.
- Yang, Z. 1994. Maximum likelihood phylogenetic estimation from DNA sequences with variable rates over sites: approximate methods. *Journal of Molecular Evolution* 39(3):306-314.
- Yang, Z. 1995. A space-time process model for the evolution of DNA sequences. *Genetics* 139(2):993-1005.
- Yang, Z. 1997. *PAML*: a program package for phylogenetic analysis by maximum likelihood. *Computer Applications in the Biosciences* 13:555-556.
- Yee, D. P., and Dill, K. A. 1993. Families and the structural relatedness among globular proteins. *Protein Science* 2:884-899.
- Yoshinaka, R., Suzuki, T., and Sato, M. 1983. Purification and some properties of anionic trypsin from the catfish pancreas. *Bulletin of the Japanese Society of Scientific Fisheries* 49(2):207-212.
- Zhang, Z., Pearson, W. R., and Miller, W. 1997. Aligning a DNA sequence with a protein sequence. *Journal of Computational Biology* 4(3):339-349.
- Zimmer, E. A., Martin, S. L., Beverley, S. M., Kan, Y. W., and Wilson, A. C. 1980. Rapid duplication and loss of genes coding for the alpha chains of hemoglobin. *Proceedings of the National Academy of Sciences* 77:2158-2162.
- Zwilling, R., and Neurath, H. 1981. Invertebrate proteases. *Methods in Enzymology* 80:633-664.