

Bibliography

- [1] **The Nobel Prize in Chemistry 2002 - Information for the Public** 2002, [http://nobelprize.org/nobel_prizes/chemistry/laureates/2002/public.html]. [The Royal Swedish Academy of Science]. (document), 2.3, 2.7
- [2] Nesvizhskii AI, Aebersold R: **Analysis, statistical validation and dissemination of large-scale proteomics datasets generated by tandem MS**. *Drug Discov Today* 2004, **9**(4):173–81. (document), 2.6
- [3] Lange E, Gropl C, Reinert K, Kohlbacher O, Hildebrandt A: **High Accuracy Peak-Picking of Proteomics Data using Wavelet Techniques**. In *Proceedings of the 11th Pacific Symposium on Biocomputing (PSB-06)* 2006:243–254. (document), 2.3, 2.5.1, 2.10, 2.5.1
- [4] **Mascot** 2005, [<http://www.matrixscience.com>]. (document), 2.5.2.1, 3.4
- [5] Tabb DL, MacCoss MJ, Wu CC, Anderson SD, Yates JRr: **Similarity among tandem mass spectra from proteomic experiments: detection, significance, and utility**. *Anal Chem* 2003, **75**(10):2470–7. (document), 1, 1, 1, 3.1, 5.1.1, 5.1.2, 5.2.7, 5.3.4, 5.4, 5.3.13
- [6] Fenyo D: **Identifying the proteome: software tools**. *Current Opinion in Biotechnology* 2000, **11**:391–395. 1, 3.1, 4.1, 5.1
- [7] Gevaert K, Vandekerckhove J: **Protein identification methods in proteomics**. *Electrophoresis* 2000, **21**(6):1145–54. 1, 4.1

- [8] Griffin TJ, Aebersold R: **Advances in proteome analysis by mass spectrometry.** *J Biol Chem* 2001, **276**:45497–500. [1](#), [3.1](#), [4.1](#), [5.1](#)
- [9] Patterson SD: **Data analysis—the Achilles heel of proteomics.** *Nat Biotechnol* 2003, **21**(3):221–2. [1](#), [2.6](#), [3.1](#), [4.1](#), [5.1](#)
- [10] Aebersold R, Mann M: **Mass spectrometry-based proteomics.** *Nature* 2003, **422**(6928):198–207. [1](#), [3.1](#), [4.1](#), [5.1](#)
- [11] Kaltschmidt E, Wittmann HG: **Ribosomal proteins. XII. Number of proteins in small and large ribosomal subunits of Escherichia coli as determined by two-dimensional gel electrophoresis.** *Proc Natl Acad Sci U S A* 1970, **67**(3):1276–82. [1](#), [4.1](#)
- [12] O’Farrell PH: **High resolution two-dimensional electrophoresis of proteins.** *J Biol Chem* 1975, **250**(10):4007–21. [1](#), [4.1](#)
- [13] Klose J, Kobalz U: **Two-dimensional electrophoresis of proteins: an updated protocol and implications for a functional analysis of the genome.** *Electrophoresis* 1995, **16**(6):1034–59. [1](#), [3.2.1](#), [4.1](#), [4.2.1](#), [5.2.1.1](#)
- [14] Blackstock W, Weir M: **Proteomics: quantitative and physical mapping of cellular proteins.** *Trends Biotech* 1999, **17**:121–127. [1](#), [4.1](#)
- [15] Quadroni M, James P: **Proteomics and automation.** *Electrophoresis* 1999, **20**:664–677. [1](#), [4.1](#)
- [16] Nordhoff E, Egelhofer V, Giavalisco P, Eickhoff H, Horn M, Przewieslik T, Theiss D, Schneider U, Lehrach H, Gobom J: **Large-gel two-dimensional electrophoresis-matrix assisted laser desorption/ionization-time of flight-mass spectrometry: an analytical challenge for studying complex protein mixtures.** *Electrophoresis* 2001, **22**(14):2844–2855. [(eng)]. [1](#), [4.1](#)
- [17] Lottspeich F, Eckerskorn C: **Internal amino acid sequence analysis of proteins separated by gel electrophoresis after tryptic digestion in polyacrylamide matrix.** *Chromatographia* 1989, :92–94. [1](#), [4.1](#)

- [18] Mann M, Hojrup P, Roepstorff P: **Use of mass spectrometric molecular weight information to identify proteins in sequence databases.** *Biol Mass Spectrom* 1993, **22**(6):338–345. [1](#), [2.3](#), [3.1](#), [4.1](#), [5.1](#)
- [19] Pappin DJC, Hojrup P, Bleasby AJ: **Rapid identification of proteins by peptide-mass fingerprinting.** *Curr. Biol.* 1993, **3**:327–332. [1](#), [3.1](#), [4.1](#)
- [20] Figgeys D, Aebersold R: **High sensitivity analysis of proteins and peptides by capillary electrophoresis-tandem mass spectrometry: recent developments in technology and applications.** *Electrophoresis* 1998, **19**(6):885–92. [1](#)
- [21] Frenz J, Bourell J, Hancock WS: **High-performance displacement chromatography-mass spectrometry of tryptic peptides of recombinant human growth hormone.** *J Chromatogr* 1990, **512**:299–314. [1](#)
- [22] Davis MT, Beierle J, Bures ET, McGinley MD, Mort J, Robinson JH, Spahr CS, Yu W, Luethy R, Patterson SD: **Automated LC-LC-MS-MS platform using binary ion-exchange and gradient reversed-phase chromatography for improved proteomic analyses.** *J Chromatogr B Biomed Sci Appl* 2001, **752**(2):281–91. [1](#)
- [23] Opitck GJ, Ramirez SM, Jorgenson JW, Moseley MAr: **Comprehensive two-dimensional high-performance liquid chromatography for the isolation of overexpressed proteins and proteome mapping.** *Anal Biochem* 1998, **258**(2):349–61. [1](#)
- [24] Johnson R, Martin S, Biemann K, Stults J, Watson J: **Novel Fragmentation Process of Peptides by Collision-Induced Decomposition in a Tandem Mass Spectrometer: Differentiation of Leucine and Isoleucine.** *Analytical Chemistry* 1987, **59**(21):2621–2625. [1](#), [3.1](#), [5.1](#)
- [25] Smith RD, Loo JA, Edmonds CG, Barinaga CJ, Udseth HR: **New developments in biochemical mass spectrometry: electrospray ionization.** *Anal Chem* 1990, **62**(9):882–99. [1](#), [3.1](#), [5.1](#)

- [26] Yates JRr: **Mass spectrometry and the age of the proteome.** *J Mass Spectrom* 1998, **33**:1–19. [1](#)
- [27] Pappin J, Hojrup P, Bleasby A: **Rapid Identification of Proteins by Peptide-Mass Fingerprinting.** *Current Biology* 1993, **3**:327–332. [1](#), [2.3](#), [2.5.2.1](#), [4.1](#)
- [28] Mann M, Wilm M: **Error-tolerant identification of peptides in sequence databases by peptide sequence tags.** *Anal Chem* 1994, **66**(24):4390–9. [1](#), [2.5.2.2](#), [5.1.1](#)
- [29] Zhang W, Chait BT: **ProFound: an expert system for protein identification using mass spectrometric peptide mapping information.** *Anal Chem* 2000, **72**(11):2482–2489. [1](#), [2.5.2.1](#), [4.1](#)
- [30] Wool A, Smilansky Z: **Precalibration of matrix-assisted laser desorption/ionization-time of flight spectra for peptide mass fingerprinting.** *Proteomics* 2002, **2**(10):1365–1373. [1](#), [1](#), [2.5.2.1](#), [3.1](#), [3.1](#), [3.1](#), [3.2.6](#), [3.3.1](#), [3.3.5](#), [3.3.6.1](#), [3.3.6.1](#), [3.3.6.2](#), [4.1](#), [4.1.1](#), [4.3.1](#), [5.1](#), [5.2.1.1](#), [5.2.3](#)
- [31] Eriksson J, Fenyo D: **A Model of random mass-matching and its use for automated significance testing in mass spectrometric proteome analysis.** *Proteomics* 2002, **2**(3):262–70. [1](#), [4.1](#)
- [32] Wolski WE, Farrow M, Emde AK, Lehrach H, Lalowski M, Reinert K: **Analytical model of peptide mass cluster centres with applications.** *Proteome Sci* 2006, **4**:18. [1](#), [2.5.1](#), [4.2.8](#)
- [33] Parker KC: **Scoring methods in MALDI peptide mass fingerprinting: ChemScore, and the ChemApplex program.** *J Am Soc Mass Spectrom* 2002, **13**:22–39. [1](#), [4.1](#)
- [34] Tabb DL, Huang Y, Wysocki VH, Yates JRr: **Influence of basic residue content on fragment ion peak intensities in low-energy collision-induced dissociation spectra of peptides.** *Anal Chem* 2004, **76**(5):1243–8. [1](#), [4.1](#)

- [35] Gras R, Muller M, Gasteiger E, Gay S, Binz PA, Bienvenut W, Hoogland C, Sanchez JC, Bairoch A, Hochstrasser DF, Appel RD: **Improving protein identification from peptide mass fingerprinting through a parameterized multi-level scoring algorithm and an optimized peak detection.** *Electrophoresis* 1999, **20**(18):3535–3550. [(eng)]. [1](#), [1](#), [3.1](#), [3.1](#), [3.3.6.1](#), [3.3.6.1](#), [4.1](#), [4.1.1](#), [4.2.3](#), [5.1](#)
- [36] Sadygow RG, Cociorva D, Yates JR: **Large-scale database searching using tandem mass spectra: Looking up the answer in the back of the book.** *Nature methods* 2004, **1**(3):195–202. [1](#), [2.5.3](#), [5.1](#), [5.2.1.2](#)
- [37] Gentzel M, Kocher T, Ponnusamy S, Wilm M: **Preprocessing of tandem mass spectrometric data to support automatic protein identification.** *Proteomics* 2003, **3**(8):1597–610. [1](#), [1](#), [3.1](#), [5.1](#), [5.1.1](#)
- [38] Tabb DL, Saraf A, Yates JR: **GutenTag: high-throughput sequence tagging via an empirically derived fragmentation model.** *Anal Chem* 2003, **75**(23):6415–21. [1](#), [2.5.2.2](#), [5.1.1](#)
- [39] Pevzner PA, Dancik V, Tang CL: **Mutation-Tolerant Protein Identification by Mass Spectrometry.** *Journal of Computational Biology* 2000, **7**(6):777–787. [1](#), [4.1](#), [5.1.1](#)
- [40] Gobom J, Mueller M, Egelhofer V, Theiss D, Lehrach H, Nordhoff E: **A calibration method that simplifies and improves accurate determination of peptide molecular masses by MALDI-TOF MS.** *Anal Chem* 2002, **74**(15):3915–3923. [(eng)]. [1](#), [1](#), [3.1](#), [3.2.2](#), [4.1.1](#), [4.2.9](#), [4.3.3](#), [5.1](#)
- [41] Wolski WE, Lalowski M, Jungblut P, Reinert K: **Calibration of mass spectrometric peptide mass fingerprint data without specific external or internal calibrants.** *BMC Bioinformatics* 2005, **6**:203, [<http://www.biomedcentral.com/1471-2105/6/203>]. [1](#), [1](#), [2.3](#), [2.5.1](#), [3.1](#), [3.2.2](#), [3.3.6.1](#), [3.3.6.2](#), [5.2.1.1](#), [6](#)

- [42] Levander F, Rognvaldsson T, Samuelsson J, James P: **Automated methods for improved protein identification by peptide mass fingerprinting.** *Proteomics* 2004, 4(9):2594–601. 1, 1, 3.1, 4.1.1, 4.2.4, 4.2.6, 4.3.1, 4.3.5, 5.1
- [43] Chamrad DC, Koerting G, Gobom J, Thiele H, Klose J, Meyer HE, Blueggel M: **Interpretation of mass spectrometry data for high-throughput proteomics.** *Anal Bioanal Chem* 2003, 376(7):1014–22. 1, 1, 3.1, 3.3.6.1, 4.1.1, 4.1.2, 4.2.6, 4.3.1, 4.3.5, 5.1, 5.2.1.1
- [44] Moore RE, Young MK, Lee TD: **Method for screening peptide fragment ion mass spectra prior to database searching.** *J Am Soc Mass Spectrom* 2000, 11(5):422–6. 1, 5.1
- [45] Sun W, Li F, Wang J, Zheng D, Gao Y: **AMASS: Software for Automatically Validating the Quality of MS/MS Spectrum from SEQUEST Results.** *Mol Cell Proteomics* 2004, 3(12):1194–9. 1, 5.1, 5.4, 6
- [46] Anderson DC, Li W, Payan DG, Noble WS: **A new algorithm for the evaluation of shotgun peptide sequencing in proteomics: support vector machine classification of peptide MS/MS spectra and SEQUEST scores.** *J Proteome Res* 2003, 2(2):137–46. 1, 2.5.3, 5.1
- [47] Keller A, Nesvizhskii AI, Kolker E, Aebersold R: **Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search.** *Anal Chem* 2002, 74(20):5383–92. 1, 2.5.3, 5.1
- [48] Yates JR, Morgan SF, Gatlin CL, Griffin PR, Eng JK: **Method To Compare Collision-Induced Dissociation Spectra of Peptides: Potential for Library Searching and Subtractive Analysis.** *Anal. Chem.* 1998, 70:3557–3565. 1, 1, 5.1.1
- [49] Beer I, Barnea E, Ziv T, Admon A: **Improving large-scale proteomics by clustering of mass spectrometry data.** *Proteomics* 2004, 4(4):950–60. 1, 1, 5.1.1

- [50] Wolski WE, Lalowski M, Martus P, Herwig R, Giavalisco P, Gobom J, Sickmann A, Lehrach H, Reinert K: **Transformation and other factors of the peptide mass spectrometry pairwise peak-list comparison process.** *BMC Bioinformatics* 2005, **6**:285. [1](#), [1](#), [1](#), [2.5.1](#), [3.1](#), [4.1.1](#)
- [51] Gay S, Binz PA, Hochstrasser DF, Appel RD: **Modeling peptide mass fingerprinting data using the atomic composition of peptides.** *Electrophoresis* 1999, **20**(18):3527–3534. [(eng)]. [1](#), [3.1](#), [3.1](#), [4.2.8](#), [4.3.1](#), [5.2.1.1](#)
- [52] Zhen Y, Xu N, Richardson B, Becklin R, Savage JR, Blake K, Peltier JM: **Development of an LC-MALDI method for the analysis of protein complexes.** *J Am Soc Mass Spectrom* 2004, **15**(6):803–22. [1](#), [2.3](#), [3.1](#)
- [53] Schmidt F, Schmid M, Jungblut PR, Mattow J, Facius A, Pleissner KP: **Iterative data analysis is the key for exhaustive analysis of peptide mass fingerprints from proteins separated by two-dimensional electrophoresis.** *J Am Soc Mass Spectrom* 2003, **14**(9):943–56. [1](#), [3.1](#), [3.1](#), [4.3.2](#), [5.3.1](#)
- [54] Cagney G, Amiri S, Premawaradena T, Lindo M, Emili A: **In silico proteome analysis to facilitate proteomics experiments using mass spectrometry.** *Proteome Science* 2003, **1**:5, [<http://www.Proteomesci.com/content/1/1/5>]. [1](#), [3.1](#)
- [55] Perkins DN, Pappin DJ, Creasy DM, Cottrell JS: **Probability-based protein identification by searching sequence databases using mass spectrometry data.** *Electrophoresis* 1999, **20**(18):3551–3567. [1](#), [2.5.2.1](#), [3.1](#), [3.1](#), [3.2.2](#), [3.3.6.2](#), [4.2.1](#), [4.3.5](#), [4.3.5](#), [4.3.6](#), [5.2.1.1](#)
- [56] Egelhofer V, Gobom J, Seitz H, Giavalisco P, Lehrach H, Nordhoff E: **Protein identification by MALDI-TOF-MS peptide mapping: A new strategy.** *Analytical Chemistry* 2002, **74**(8):1760–1771. [1](#), [4.1.1](#), [5.2.3](#)
- [57] Schuerenberg M, Luebbert C, Eickhoff H, Kalkum M, Lehrach H, Nordhoff E: **Prestructured MALDI-MS sample supports.** *Anal Chem* 2000, **72**(15):3436–42. [1](#), [4.1.1](#), [4.1.2](#)

- [58] Guilhaus M: **Principles and Instrumentation in Time-of-Flight Mass Spectrometry**. *JOURNAL OF MASS SPECTROMETRY* 1995, **30**:1519–1532. [1](#), [2.3](#), [4.1](#), [4.1.1](#)
- [59] Bookstein F: **Principal Warps: Thin-Plate Splines and the Decomposition of Deformations**. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 1989, **11**(6):567–585. [1](#), [4.1.1](#), [4.2.7](#), [4.2.7](#), [4.3.3](#)
- [60] Alfassi ZB: **On the normalization of a mass spectrum for comparison of two spectra**. *J Am Soc Mass Spectrom* 2004, **15**(3):385–387. [1](#), [5.1.2](#), [5.2.6](#), [5.2.7](#)
- [61] Rasmussen GT, Isenhour TL: **The Evaluation of Mass Spectral Search Algorithms**. *Journal of Chemical Information and Computer Sciences* 1979, **19**(3):179 – 186. [1](#), [5.1.2](#), [5.2.6](#)
- [62] Herwig R, Poustka AJ, Müller C, Lehrach H, O'Brien J: **Large-scale Clustering of cDNA Fingerprinting Data**. *Genome Res.* 1999, **9**:1093–1105. [1](#), [5.1.2](#), [5.2.5](#)
- [63] Monson M, Field F: **Chemical ionisation mass spectrometry. I. General introduction**. *J. Am. Chem. Soc.* 1966, **88**:2621. [2.2](#)
- [64] MacFarlane R, Torgerson D: **Californium-252 plasma desorption mass spectroscopy**. *Science* 1976, **191**:920–925. [2.2](#)
- [65] Barber M, Bordoli RS, Sedgwick RD, Tyler AN, Bycroft BW: **Fast atom bombardment mass spectrometry of bleomycin A2 and B2 and their metal complexes**. *Biochem Biophys Res Commun* 1981, **101**(2):632–8. [2.2](#)
- [66] Klose J: **Protein mapping by combined isoelectric focusing and electrophoresis of mouse tissues. A novel approach to testing for induced point mutations in mammals**. *Humangenetik* 1975, **26**(3):231–43. [2.3](#)

- [67] Li DQ, Wang L, Fei F, Hou YF, Luo JM, Zeng R, Wu J, Lu JS, Di GH, Ou ZL, Xia QC, Shen ZZ, Shao ZM: **Identification of breast cancer metastasis-associated proteins in an isogenic tumor metastasis model using two-dimensional gel electrophoresis and liquid chromatography-ion trap-mass spectrometry.** *Proteomics* 2006, **6**(11):3352–68. [2.3](#)
- [68] Steel LF, Haab BB, Hanash SM: **Methods of comparative proteomic profiling for disease diagnostics.** *J Chromatogr B Analyt Technol Biomed Life Sci* 2005, **815**(1-2):275–84. [2.3](#)
- [69] Vandahl BSB, Birkelund S, Christiansen G: **Genome and proteome analysis of Chlamydia.** *Proteomics* 2004, **4**(10):2831–42. [2.3](#)
- [70] Gobom J, Schürenberg M, Mueller M, Theiss D, Lehrach H, Nordhoff E: **Alpha-cyano-4-hydroxycinnamic acid affinity sample preparation. A protocol for MALDI-MS peptide analysis in proteomics.** *Analytical Chemistry* 2001, **73**(3):434–438. [2.3](#), [3.1](#), [4.2.9](#)
- [71] Gustavsson N, Greber B, Kreitler T, Himmelbauer H, Lehrach H, Gobom J: **A proteomic method for the analysis of changes in protein concentrations in response to systemic perturbations using metabolic incorporation of stable isotopes and mass spectrometry.** *Proteomics* 2005, **5**(14):3563–70. [2.3](#)
- [72] Dole M, Mach L, Hines R, Mobley R, Ferguson L, Alice M: **Molecular beams of macroions.** *J. Chem. Phys.* 1968, **49**:2240–2247. [2.4](#)
- [73] Fenn JB, Mann M, Meng CK, Wong SF, Whitehouse CM: **Electrospray ionization for mass spectrometry of large biomolecules.** *Science* 1989, **246**(4926):64–71. [2.4](#), [4.1](#)
- [74] Aleksandrov M, Gall L, Krasnov V, Nikolaev V, Pavlenko V, Shkurov V: *Dokl Akad Nauk SSSR.* 1984, **277**:379–383. [2.4](#)

- [75] Hunt DF, Yates JRr, Shabanowitz J, Winston S, Hauer CR: **Protein sequencing by tandem mass spectrometry**. *Proc Natl Acad Sci U S A* 1986, **83**(17):6233–7. [2.4](#)
- [76] Stults JT, Watson JT: **Identification of a new type of fragment ion in the collisional activation spectra of peptides allows leucine/isoleucine differentiation**. *Biomed Environ Mass Spectrom* 1987, **14**(10):583–6. [2.4](#)
- [77] Kreitler T: **Oral Communication** 2003. [2.5.1](#), [4.2.4](#)
- [78] Tan HW, Brown SD: **Wavelet analysis applied to removing non-constant, varying spectroscopic background in multivariate calibration**. *Journal of Chemometrics* 2002, **16**(5):228 – 240. [2.5.1](#)
- [79] **Wikipedia is a Web-based, free-content encyclopedia** 2004, [<http://www.wikipedia.org>]. [2.5.1](#), [3.1](#)
- [80] Breen EJ, Hopwood FG, Williams KL, Wilkins MR: **Automatic poisson peak harvesting for high throughput protein identification**. *Electrophoresis* 2000, **21**(11):2243–51. [2.5.1](#)
- [81] Yasui Y, McLerran D, Adam BL, Winget M, Thornquist M, Feng Z: **An Automated Peak Identification/Calibration Procedure for High-Dimensional Protein Measures From Mass Spectrometers**. *J Biomed Biotechnol* 2003, **2003**(4):242–248. [2.5.1](#)
- [82] Strittmatter EF, Rodriguez N, Smith RD: **High mass measurement accuracy determination for proteomics using multivariate regression fitting: application to electrospray ionization time-of-flight mass spectrometry**. *Anal Chem* 2003, **75**(3):460–8. [2.5.1](#), [4.1](#), [5.1](#)
- [83] Kempka M, Sjobahl J, Bjork A, Roeraade J: **Improved method for peak picking in matrix-assisted laser desorption/ionization time-of-flight mass spectrometry**. *Rapid Commun Mass Spectrom* 2004, **18**(11):1208–12. [2.5.1](#)

- [84] **Bruker Daltonics - enabling life science tools based on mass spectrometry.** 2004, [<http://www.bdal.com>]. 2.5.1, 1, 2, 3, 4.1.2, 1, 2, 3, 4.3.6
- [85] Karlin S, Altschul SF: **Methods for assessing the statistical significance of molecular sequence features by using general scoring schemes.** *Proc Natl Acad Sci U S A* 1990, **87**(6):2264–8. 2.5.2.1
- [86] Kuster B, Mortensen P, Andersen JS, Mann M: **Mass spectrometry allows direct identification of proteins in large genomes.** *Proteomics* 2001, **1**(5):641–50. 2.5.2.2
- [87] Choudhary JS, Blackstock WP, Creasy DM, Cottrell JS: **Interrogating the human genome using uninterpreted mass spectrometry data.** *Proteomics* 2001, **1**(5):651–67. 2.5.2.2
- [88] Tabb DL, Eng JK, Yates JRr: **Protein Identification by SEQUEST** 2001. 2.5.2.2
- [89] Bafna V, Edwards N: **SCOPE: a probabilistic model for scoring tandem mass spectra against a peptide database.** *BIOINFORMATICS* 2001, **17**:13–29. 2.5.2.2
- [90] Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ: **Basic local alignment search tool.** *J Mol Biol* 1990, **215**(3):403–10. 2.5.2.2
- [91] Liska AJ, Sunyaev S, Shilov IN, Schaeffer DA, Shevchenko A: **Error-tolerant EST database searches by tandem mass spectrometry and multiTag software.** *Proteomics* 2005, **5**(16):4118–22. 2.5.2.2
- [92] Taylor JA, Johnson RS: **Implementation and uses of automated de novo peptide sequencing by tandem mass spectrometry.** *Anal Chem* 2001, **73**(11):2594–604. 2.5.2.2
- [93] Dancik V, Addona TA, Clauser KR, Vath JE, Pevzner PA: **De novo peptide sequencing via tandem mass spectrometry.** *J Comput Biol* 1999, **6**(3-4):327–42. 2.5.2.2

- [94] de Cossio JF, Gonzalez J, Satomi Y, Shima T, Okumura N, Besada V, Betancourt L, Padron G, Shimonishi Y, Takao T: **Automated interpretation of low-energy collision-induced dissociation spectra by SeqMS, a software aid for de novo sequencing by tandem mass spectrometry.** *Electrophoresis* 2000, **21**(9):1694–9. [2.5.2.2](#)
- [95] Chen T, Kao M, Tepel M, Rush J, Church G: **A dynamic programming approach to de novo peptide sequencing via tandem mass spectrometry.** *Journal of Computational Biology* 2001, **8**(6):571–583. [2.5.2.2](#)
- [96] Cannon WR, Jarman KD: **Improved peptide sequencing using isotope information inherent in tandem mass spectra.** *Rapid Commun Mass Spectrom* 2003, **17**(15):1793–801. [2.5.2.2](#)
- [97] Molecular, Proteomics C: **PUBLICATION GUIDELINES FOR THE ANALYSIS AND DOCUMENTATION OF PEPTIDE AND PROTEIN IDENTIFICATIONS** 2006, [http://www.mcponline.org/misc/ParisReport_Final.shtml]. [2.5.3](#), [2.6](#)
- [98] Chamrad D, Meyer HE: **Valid data from large-scale proteomics studies.** *Nat Methods* 2005, **2**(9):647–8. [2.5.3](#)
- [99] Han DK, Eng J, Zhou H, Aebersold R: **Quantitative profiling of differentiation-induced microsomal proteins using isotope-coded affinity tags and mass spectrometry.** *Nat Biotechnol* 2001, **19**(10):946–51. [2.5.3](#)
- [100] Tabb DL, McDonald WH, Yates JRr: **DTASelect and Contrast: tools for assembling and comparing protein identifications from shotgun proteomics.** *J Proteome Res* 2002, **1**:21–6. [2.5.3](#)
- [101] Eddes JS, Kapp EA, Frecklington DF, Connolly LM, Layton MJ, Moritz RL, Simpson RJ: **CHOMPER: a bioinformatic tool for rapid validation of tandem mass spectrometry search results associated with high-throughput proteomic strategies.** *Proteomics* 2002, **2**(9):1097–103. [2.5.3](#)

- [102] Petritis K, Kangas LJ, Ferguson PL, Anderson GA, Pasa-Tolic L, Lipton MS, Auberry KJ, Strittmatter EF, Shen Y, Zhao R, Smith RD: **Use of artificial neural networks for the accurate prediction of peptide liquid chromatography elution times in proteome analyses.** *Anal Chem* 2003, **75**(5):1039–48. [2.5.3](#)
- [103] Cargile BJ, Bundy JL, Freeman TW, Stephenson JLJ: **Gel based isoelectric focusing of peptides and the utility of isoelectric point in protein identification.** *J Proteome Res* 2004, **3**:112–9. [2.5.3](#)
- [104] Apweiler R, Bairoch A, Wu CH: **Protein sequence databases.** *Curr Opin Chem Biol* 2004, **8**:76–80. [3.1](#), [4.1](#)
- [105] Giles J: **Internet encyclopaedias go head to head.** *Nature* 2005, **438**(7070):900–1. [3.1](#)
- [106] Giavalisco P, Nordhoff E, Kreitler T, Kloppel KD, Lehrach H, Klose J, Gobom J: **Proteome analysis of Arabidopsis thaliana by two-dimensional gel electrophoresis and matrix-assisted laser desorption/ionisation-time of flight mass spectrometry.** *Proteomics* 2005, **5**(7):1902–13. [3](#), [3.2.1](#), [4.1.2](#), [3](#), [4.2.1](#), [5.2.1.1](#)
- [107] Pruitt KD, Tatusova T, Maglott DR: **NCBI Reference Sequence project: update and current status.** *Nucleic Acids Res* 2003, **31**:34–7. [3.2.3](#), [4.1.2](#)
- [108] Glockner FO, Kube M, Bauer M, Teeling H, Lombardot T, Ludwig W, Gade D, Beck A, Borzym K, Heitmann K, Rabus R, Schlesner H, Amann R, Reinhardt R: **Complete genome sequence of the marine planctomycete Pirellula sp. strain 1.** *Proc Natl Acad Sci U S A* 2003, **100**(14):8298–303. [3.2.3](#), [4.1.2](#)
- [109] Bairoch A, Apweiler R, Wu CH, Barker WC, Boeckmann B, Ferro S, Gasteiger E, Huang H, Lopez R, Magrane M, Martin MJ, Natale DA, O'Donovan C, Redaschi N, Yeh LSL: **The Universal Protein Resource (UniProt).** *Nucleic Acids Res* 2005, **33**(Database issue):D154–9. [3.2.3](#), [3.2.5](#), [3.3.1](#), [3.3.1](#)

- [110] Emde AK: **Protein Sequence Digester** 2004, [<http://www.inf.fu-berlin.de/~emde/>]. 3.2.4
- [111] Becker RA, Chambers JM, Wilks AR: *The New S Language*. London: Chapman & Hall 1988. 3.3.3
- [112] Samuelsson J, Dalevi D, Levander F, Rognvaldsson T: **Modular, Scriptable, and Automated Analysis Tools for High-Throughput Peptide Mass Fingerprinting**. *Bioinformatics* 2004. 3.3.6.1, 4.1, 4.1.1, 4.1.2, 4.3.1
- [113] Venables WN, Ripley BD: *Modern Applied Statistics with S. Fourth Edition*. Springer 2002, [<http://www.stats.ox.ac.uk/pub/MASS4/>]. [ISBN 0-387-95457-0]. 3.3.6.1, 4.2.8
- [114] Tanaka K, Waki H, Ido Y, Akita S, Yoshida Y, Yoshida T, Matsuo T: **Protein and polymer analyses up to m/z 100 000 by laser ionization time-of-flight mass spectrometry**. *Rapid Communications in Mass Spectrometry* 1988;, 2(8):151–153. 4.1
- [115] Karas M, Hillenkamp F: **Laser Desorption Ionization of proteins with molecular masses exceeding 10 000 daltons**. *Anal Chem* 1988, 60:2299–2301. 4.1
- [116] Colby SM, King TB, Reilly JP: **Improving the Resolution of MALDI TOF Mass Spectrometry by Exploiting the Correlation Between Ion Position and Velocity**. *Rapid Comm. Mass Spectrom.* 1994, 8:865–868. 4.1
- [117] Whittall RM, Li L: **High-resolution matrix-assisted laser desorption/ionization in a linear time-of-flight mass spectrometer**. *Anal Chem* 1995, 67(13):1950–4. 4.1
- [118] Brown RS, Lennon JJ: **Mass resolution improvement by incorporation of pulsed ion extraction in a matrix-assisted laser desorption/ionization linear time-of-flight mass spectrometer**. *Anal Chem* 1995, 67(13):1998–2003. 4.1

- [119] Takach EJ, Hines WM, Patterson DH, Juhasz P, Falick AM, Vestal ML, Martin SA: **Accurate mass measurements using MALDI-TOF with delayed extraction.** *J Protein Chem* 1997, **16**(5):363–9. [4.1](#)
- [120] Bantscheff M, Duempelfeld B, Kuster B: **An improved two-step calibration method for matrix-assisted laser desorption/ionization time-of-flight mass spectra for proteomics.** *Rapid Commun Mass Spectrom* 2002, **16**(19):1892–5. [4.1.1](#), [4.2.9](#), [4.3.4](#)
- [121] Moskovets E, Chen HS, Pashkova A, Rejtar T, Andreev V, Karger BL: **Closely spaced external standard: a universal method of achieving 5 ppm mass accuracy over the entire MALDI plate in axial matrix-assisted laser desorption/ionization time-of-flight mass spectrometry.** *Rapid Commun Mass Spectrom* 2003, **17**(19):2177–87. [4.1.1](#), [4.3.3](#)
- [122] Gobom J, Mueller M, Egelhofer V, Theiss D, Lehrach H, Nordhoff E: **A Calibration Method that Simplifies and Improves Accurate Determination of Peptide Molecular Masses by MALDI-TOF-MS.** *Analytical Chemistry* 2002, **74**(8):3915–3923. [4.2.3](#), [4.3.4](#)
- [123] Lee K, Bae D, Lim D: **Evaluation of parameters in peptide mass fingerprinting for protein identification by MALDI-TOF mass spectrometry.** *Mol Cells* 2002, **13**(2):175–84. [4.2.3](#)
- [124] Härdle W, Müller M, Sperlich S, Werwatz A: *Nonparametric and Semiparametric Models - An Introduction*. Springer, Heidelberg 2004, [<http://www.quantlet.com/mdstat/scripts/spm/html/spmhtml.html>]. [4.2.4](#)
- [125] Donato G, Belongie S: **Approximation Methods for Thin Plate Spline Mappings and Principal Warps.** In *Computer Vision - ECCV 2002: 7th European Conference on Computer Vision, Copenhagen, Denmark, May 28–31, 2002. Proceedings, Part III*, Lecture Notes in Computer Science. Edited by Heyden A, Sparr G, Nielsen M, Johansen P, Springer-Verlag Heidelberg 2002:21–31. [4.2.7](#), [4.2.7](#)

- [126] Green P, Silverman B: *Nonparametric Regression and Generalized Linear Models: A Roughness Penalty Approach*. Chapman and Hall 1994. 4.2.7
- [127] Nychka D: **fields - A collection of programs based in [R,S] for curve and function fitting with an emphasis on spatial data**. 2004, [<http://www.cgd.ucar.edu/stats/Software/Fields/>]. 4.2.7, 4.3.3
- [128] Chambers JM: **Linear models**. In *Statistical Models in S.* Edited by Chambers J, Hastie T, Wadsworth & Brooks/Cole 1992. 4.2.8
- [129] Chambers JM, Hastie TJ: *Statistical Models in S*. London: Chapman & Hall 1992. 4.2.9
- [130] Hastie T, Tibshirani R: *Generalized Additive Models*. Chapman and Hall 1990. 4.2.9
- [131] Cleveland W, Grosse E, Shyu W: **Local Regression Models**. In *Statistical Models in S.* Edited by Chambers J, Hastie T, Wadsworth & Brooks/Cole 1992. 4.2.9
- [132] Schrijver A: *Combinatorial Optimization - Polyhedra and Efficiency*. Berlin: Springer-Verlag 2003. 4.2.11, 4.3.2
- [133] Thiede B, Lamer S, Mattow J, Siejak F, Dimmler C, Rudel T, Jungblut PR: **Analysis of missed cleavage sites, tryptophan oxidation and N-terminal pyroglutamylation after in-gel tryptic digestion**. *Rapid Commun Mass Spectrom* 2000, 14(6):496–502. 4.3.1
- [134] Härdle W, Simar L: *Applied Multivariate Statistical Analysis*. Springer, Heidelberg 2003, [<http://www.quantlet.com/mdstat/scripts/mva/htmlbook/mvahtml.html>]. 4.3.2, 5.2.7
- [135] Handl A: *Multivariate Analysemethoden - Theorie und Praxis multivariater Verfahren unter besonderer Berücksichtigung von S-PLUS*. Springer, Heidelberg 2003, [<http://www.quantlet.com/mdstat/scripts/mst/html>]. 4.3.2
- [136] **R for Proteomics** [<http://r4proteomics.sourceforge.net>]. 4.3.5

- [137] R Development Core Team: **R: A language and environment for statistical computing**. R Foundation for Statistical Computing, Vienna, Austria 2004, [<http://www.R-project.org>]. [ISBN 3-900051-00-3]. 4.3.6, 5.1.3
- [138] Gentleman RC, Carey VJ, Bates DM, Bolstad B, Dettling M, Dudoit S, Ellis B, Gautier L, Ge Y, Gentry J, Hornik K, Hothorn T, Huber W, Iacus S, Irizarry R, Li FLC, Maechler M, Rossini AJ, Sawitzki G, Smith C, Smyth G, Tierney L, Yang JYH, Zhang J: **Bioconductor: Open software development for computational biology and bioinformatics**. *Genome Biology* 2004, 5:R80, [<http://genomebiology.com/2004/5/10/R80>]. 4.3.6, 4.3.7, 5.1.3
- [139] **Bioconductor - open source software for bioinformatics** 2004, [<http://www.bioconductor.org>]. 4.3.7, 5.1.3, 5.2.1.1
- [140] Leisch F: **Sweave and Beyond: Computations on Text Documents**. In *Proceedings of the 3rd International Workshop on Distributed Statistical Computing*. Edited by Hornik K, Leisch F, Zeileis A, Technische Universität Wien, Vienna, Austria 2003. 4.3.7
- [141] Clauser KR, Baker P, Burlingame AL: **Role of accurate mass measurement (+/- 10 ppm) in protein identification strategies employing MS or MS/MS and database searching**. *Anal Chem* 1999, 71(14):2871–82. 5.1.1
- [142] Svetnik V, Liaw AI: **Detecting Novel Samples in Mass Spectral Data: A Clustering Approach**. In *Proceedings of the 33rd Symposium on the Interface*. Edited by Wegman E, Braverman A, Goodman A, Smyth P 2001:321–328. 5.1.1, 5.2.7, 5.3.4
- [143] An Z, Harris G, Zink D, Giacobbe R, Lu P, Sangari R, Bills G, Svetnik V, Gunter B, Liaw A, Masurekar P, Liesch J, Gould S, , Strohl W: **Expression of Cosmid-Size DNA of Slow-Growing Fungi in Aspergillus Nidulans for Secondary Metabolite Screening**. In *Handbook of*

- Industrial Mycology*. Edited by An Z, New York: Marcel Dekker 2003:167–187. [5.1.1](#)
- [144] Li J, Zhang Z, Rosenzweig J, Yang Y, Chan D: **Proteomics and bioinformatics approaches for identification of serum biomarkers to detect breast cancer**. *Clinical Chemistry* 2002, **48**(8):1296–1304. [5.1.1](#)
- [145] Jeffries N: **Performance of a genetic algorithm for mass spectrometry proteomics**. *BMC Bioinformatics* 2004, **5**:180, [<http://www.biomedcentral.com/1471-2105/5/180>]. [5.1.1](#)
- [146] Adam B, Qu Y, Davis J, Ward M, Clements M, Cazares L, Semmes O, Schellhammer P, Yasui Y, Feng Z, Wright G: **Serum protein fingerprinting coupled with a pattern-matching algorithm distinguishes prostate cancer from benign prostate hyperplasia and healthy men**. *Cancer Research* 2002, **62**:3609–3614. [5.1.1](#)
- [147] Petricoin E, Ardekani A, Hitt B, Levine P, Fusaro V, Steinberg S, Mills G, Simone C, Fishman D, Kohn E, Liotta L: **Use of proteomic patterns in serum to identify ovarian cancer**. *Lancet* 2002, **359**:572–577. [5.1.1](#)
- [148] Wan KX, Vidavsky I, Gross ML: **Comparing similar spectra: from similarity index to spectral contrast angle**. *J Am Soc Mass Spectrom* 2002, **13**:85–88. [5.1.2](#), [5.2.7](#), [5.3.9](#)
- [149] Sickmann A, Reinders J, Wagner Y, Joppich C, Zahedi R, Meyer HE, Schonfish B, Perschil I, Chacinska A, Guiard B, Rehling P, Pfanner N, Meisinger C: **The proteome of *Saccharomyces cerevisiae* mitochondria**. *Proc Natl Acad Sci U S A* 2003, **100**(23):13207–12. [5.2.1.2](#)
- [150] Wagner Y, Sickmann A, Meyer HE, Daum G: **Multidimensional nano-HPLC for analysis of protein complexes**. *J Am Soc Mass Spectrom* 2003, **14**(9):1003–11. [5.2.1.2](#)
- [151] Issel-Tarver L, Christie KR, Dolinski K, Andrada R, Balakrishnan R, Ball CA, Binkley G, Dong S, Dwight SS, Fisk DG, Harris M, Schroeder M,

- Sethuraman A, Tse K, Weng S, Botstein D, Cherry JM: **Saccharomyces Genome Database**. *Methods Enzymol* 2002, **350**:329–46. [5.2.1.2](#)
- [152] Yates JR, Eng JK, McCormack AL, Schieltz D: **Method to correlate tandem mass spectra of modified peptides to amino acid sequences in the protein database**. *Anal Chem* 1995, **67**(8):1426–36. [5.2.1.2](#)
- [153] Kececioglu J: **The maximum weight trace problem in multiple sequence alignment**. In *Proceedings of the 4th Symposium on Combinatorial Pattern Matching (CPM 93)*, no. 684 in Lecture Notes in Computer Science, Springer 1993:106–119. [5.2.2](#)
- [154] Kececioglu JD, Lenhof HP, Mehlhorn K, Mutzel P, Reinert K, Vingron M: **A Polyhedral Approach to Sequence Alignment Problems**. *Discrete Applied Mathematics* 2000, **104**:143–186. [5.2.2](#)
- [155] Zhang N, Aebersold R, Schwikowski B: **ProbID: A probabilistic algorithm to identify peptides through sequence database searching using tandem mass spectral data**. *Proteomics* 2002, **2**:1406–1412. [5.2.3](#), [5.3.5](#)
- [156] Hageman JA, Wehrens R, Gelder RD, Buydens LMC: **Powder Pattern Indexing Using the Weighted Crosscorrelation and Genetic Algorithms**. *Journal of Computational Chemistry* 2003, **24**(9):1043–1051. [5.2.3](#), [5.3.5](#)
- [157] Gower JC, Legendre P: **Metric and Euclidean Properties of Dissimilarity Coefficients**. *Journal of classification*. 1986, :5–48. [5.2.5](#), [5.3.2](#)
- [158] Fowlkes EB, Mallows CL: **A method for comparing two hierarchical clusterings**. *J. Am. Stat. Assoc.* 1983, **78**:553–569. [5.2.5](#)
- [159] Cover TM, Thomas J: *Elements of Information Theory*. New York: J.Wiley and Sons 1991. [5.2.5](#)

- [160] Herwig R (Ed): *Large-scale information theoretic clustering with application to the analysis of genetic fingerprinting data (PhD)*. Berlin: Logos 2001. [5.2.5](#)
- [161] Becker RA, Chambers JM, Wilks AR: *The New S Language*. Wadsworth & Brooks/Cole 1988. [5.2.6](#), [5.2.7](#)
- [162] Lance GN, Williams WT: **Mixed-Data Classificatory Programs I - Agglomerative Systems**. *Australian Computer Journal* 1967, 1:15–20. [5.2.7](#)
- [163] **Comprehensive R Archive Network**[\[http://cran.r-project.org\]](http://cran.r-project.org). [5.2.8](#)
- [164] Ross S: *A First Course In Probability*. Prentice Hall 2003. [5.3.1](#)
- [165] Fox J: *Applied Regression Analysis, Linear Models, and Related Methods*. Sage Publications 1997. [5.3.4](#), [5.3.7](#)
- [166] Bland JM, Altman DG: **Measuring agreement in method comparison studies**. *Stat Methods Med Res.* 1999, 8(2):135–60. [5.3.4](#)
- [167] Zhou XH, McClish DK, Obuchowski NA: *Statistical Methods in Diagnostic Medicine*. Wiley 2002. [5.3.6](#)
- [168] Santner T, Williams B, Notz W: *The Design and Analysis of Computer Experiments*. Springer Series in Statistics, Springer Verlag New York 2003. [5.4](#)