# Bibliography

[1] Rakesh Agrawal and Ramakrishnan Srikant. Fast algorithms for mining association rules. In Jorge B. Bocca, Matthias Jarke, and Carlo Zaniolo, editors, *Proc. 20th Int. Conf. Very Large Data Bases, VLDB*, pages 487–499. Morgan Kaufmann, 12–15 1994. ISBN 1-55860-153-8.

[2] G D Bader, D Betel, and C W Hogue. Bind: the biomolecular interaction network database. *Nucleic Acids Res*, 31(1):248–250, Jan 2003. URL `http://www.hubmed.org/display.cgi?uids=12519993`.

[3] GD Bader and CW Hogue. Analyzing yeast protein-protein interaction data obtained from different sources. *Nat Biotechnol*, 20(10):991–997, Oct 2002.

[4] GD Bader and CW Hogue. An automated method for finding molecular complexes in large protein interaction networks. *BMC Bioinformatics*, 4(1):2–2, Jan 2003.

[5] Béla Bollobás. *Random Graphs*. Cambridge University Press, second edition, 2001.

[6] E Bolten, A Schliep, S Schneckener, D Schomburg, and R Schrader. Clustering protein sequences–structure prediction by transitive homology. *Bioinformatics*, 17(10):935–41, Oct 2001.

[7] T. Bouwmeester, A. Bauch, H. Ruffner, PO. Angrand, G. Bergamini, K. Croughton, C. Cruciat, D. Eberhard, J. Gagneur, S. Ghidelli, C. Hopf, B. Huhse, R. Mangano, AM. Michon, M. Schirle, J. Schlegl, M. Schwab, MA. Stein, A. Bauer, G. Casari, G. Drewes, AC. Gavin, DB. Jackson, G. Joberty, G. Neubauer, J. Rick, B. Kuster, and G. Superti-Furga. A physical and functional map of the human TNF-alpha/NF-kappa B signal transduction pathway. *Nat Cell Biol*, 6(2):97–105, Feb 2004.

[8] S Brohée and J van Helden. Evaluation of clustering algorithms for protein-protein interaction networks. *BMC Bioinformatics*, 7:488–488, 2006. URL `http://www.hubmed.org/display.cgi?uids=17087821`.

[9] G Butland, JM Peregrin-Alvarez, J Li, W Yang, X Yang, V Canadien, A Starostine, D Richards, B Beattie, N Krogan, M Davey, J Parkinson, J Greenblatt, and A Emili. Interaction network containing conserved and essential protein complexes in Escherichia coli. *Nature*, 433(7025):531–7, Feb 2005.

[10] M. Deng, F. Sun, and T. Chen. Assessment of the reliability of protein-protein interactions and protein function prediction. *Pac Symp Biocomput*, pages 140–51, 2003.

[11] M. Deng, K. Zhang, S. Mehta, T. Chen, and F. Sun. Prediction of protein function using protein-protein interaction data. *J Comput Biol*, 10(6):947–60, 2003.

[12] M. B. Eisen, P. T. Spellman, P. O. Brown, and D. Botstein. Cluster analysis and display of genome-wide expression patterns. *Proc Natl Acad Sci U S A*, 95(25):14863–14868, Dec 1998.

[13] M Fellenberg, K Albermann, A Zollner, H W Mewes, and J Hani. Integrative analysis of protein interaction data. *Proc Int Conf Intell Syst Mol Biol*, 8:152–161, 2000.

[14] A C Gavin, P Aloy, P Grandi, R Krause, M Boesche, M Marzioch, C Rau, L J Jensen, S Bastuck, B Dümpelfeld, A Edelmann, M A Heurtier, V Hoffman, C Hoefert, K Klein, M Hudak, A M Michon, M Schelder, M Schirle, M Remor, T Rudi, S Hooper, A Bauer, T Bouwmeester, G Casari, G Drewes, G Neubauer, J M Rick, B Kuster, P Bork, R B Russell, and Superti-Furga. Proteome survey reveals modularity of the yeast cell machinery. *Nature*, Jan 2006. URL http://www.hubmed.org/display.cgi?uids=16429126.

[15] A. C. Gavin, M. Bosche, R. Krause, P. Grandi, M. Marzioch, A. Bauer, J. Schultz, J. M. Rick, A. M. Michon, C. M. Cruciat, M. Remor, C. Hofert, M. Schelder, M. Brajenovic, H. Ruffner, A. Merino, K. Klein, M. Hudak, D. Dickson, T. Rudi, V. Gnau, A. Bauch, S. Bastuck, B. Huhse, C. Leutwein, M. A. Heurtier, R. R. Copley, A. Edelmann, E. Querfurth, V. Rybin, G. Drewes, M. Raida, T. Bouwmeester, P. Bork, B. Seraphin, B. Kuster, G. Neubauer, and G. Superti-Furga. Functional organization of the yeast proteome by systematic analysis of protein complexes. *Nature*, 415(6868):141–147, Jan 2002.

[16] A. C. Gavin and G. Superti-Furga. Protein complexes and proteome organization from yeast to man. *Curr Opin Chem Biol*, 7(1):21–27, Feb 2003.

[17] M A Gilchrist, L A Salter, and A Wagner. A statistical framework for combining and interpreting proteomic datasets. *Bioinformatics*, 20(5):689–700, Mar 2004. URL http://www.hubmed.org/display.cgi?uids=15033876.

[18] E Hartuv and R Shamir. A clustering algorithm based on graph connectivity. *Information processing letters*, 76:175–181, 1999.

[19] H Hishigaki, K Nakai, T Ono, A Tanigami, and T Takagi. Assessment of prediction accuracy of protein function from protein–protein interaction data. *Yeast*, 18(6):523–531, Apr 2001.

[20] Y. Ho, A. Gruhler, A. Heilbut, G. D. Bader, L. Moore, S. L. Adams, A. Millar, P. Taylor, K. Bennett, K. Boutilier, L. Yang, C. Wolting, I. Donaldson, S. Schandorff, J. Shewnarane, M. Vo, J. Taggart, M. Goudreault, B. Muskat, C. Alfarano, D. Dewar, Z. Lin, K. Michalickova, A. R. Willems, H. Sassi, P. A. Nielsen, K. J. Rasmussen, J. R. Andersen, L. E. Johansen, L. H. Hansen, H. Jespersen, A. Podtelejnikov, E. Nielsen, J. Crawford, V. Poulsen, B. D. Sorensen, J. Matthiesen, R. C. Hendrickson, F. Gleeson, T. Pawson, M. F. Moran, D. Durocher, M. Mann, C. W. Hogue, D. Figeys, and M. Tyers. Systematic identification of protein complexes in saccharomyces cerevisiae by mass spectrometry. *Nature*, 415(6868):180–183, Jan 2002.

[21] J. Hollunder, A. Beyer, and T. Wilhelm. Identification and characterization of protein subcomplexes in yeast. *Proteomics*, 5(8):2082–2089, May 2005.

[22] T Ito, T Chiba, R Ozawa, M Yoshida, M Hattori, and Y Sakaki. A comprehensive two-hybrid analysis to explore the yeast protein interactome. *Proc Natl Acad Sci U S A*, 98(8):4569–4574, Apr 2001. URL http://www.hubmed.org/display.cgi?uids=11283351.

[23] P. Kemmeren, NL. van Berkum, J. Vilo, T. Bijma, R. Donders, A. Brazma, and FC. Holstege. Protein interaction verification and functional annotation by integrated analysis of genome-scale data. *Mol Cell*, 9(5):1133–1143, May 2002.

[24] R. Krause, C. von Mering, and P. Bork. A comprehensive set of protein complexes in yeast: mining large scale protein-protein interaction screens. *Bioinformatics*, 19(15):1901–1908, Oct 2003.

[25] R. Krause, C. von Mering, P. Bork, and T. Dandekar. Shared components of protein complexes–versatile building blocks or biochemical artefacts? *Bioessays*, 26(12):1333–1343, Dec 2004.

[26] N J Krogan, G Cagney, H Yu, G Zhong, X Guo, A Ignatchenko, J Li, S Pu, N Datta, A P Tikuisis, T Punna, J M Peregrín-Alvarez, M Shales, X Zhang, M Davey, M D Robinson, A Paccanaro, J E Bray, A Sheung, B Beattie, D P Richards, V Canadien, A Lalev, F Mena, P Wong, A Starostine, M M Canete, J Vlasblom, S Wu, C Orsi, S R Collins, S Chandran, R Haw, J J Rilstone, K Gandi, N J Thompson, G Musso, P St Onge, S Ghanny, M H Lam, G Butland, A M Altaf-Ul, S Kanaya, A Shilatifard, E O'Shea, J S Weissman, C J Ingles, T R Hughes, J Parkinson, M Gerstein, S J Wodak, A Emili, and J F Greenblatt. Global landscape of protein complexes in the yeast saccharomyces cerevisiae. *Nature*, 440(7084):637–643, Mar 2006. URL http://www.hubmed.org/display.cgi?uids=16554755.

[27] N. J. Krogan, W. T. Peng, G. Cagney, M. D. Robinson, R. Haw, G. Zhong, X. Guo, X. Zhang, V. Canadien, D. P. Richards, B. K. Beattie, A. Lalev, W. Zhang, A. P. Davierwala, S. Mnaimneh, A. Starostine, A. P. Tikuisis, J. Grigull, N. Datta, J. E. Bray, T. R. Hughes, A. Emili, and J. F. Greenblatt. High-definition macromolecular composition of yeast rna-processing complexes. *Mol Cell*, 13(2):225–39, Jan 2004.

[28] A. Krogh, M. Brown, I. S. Mian, K. Sjolander, and D. Haussler. Hidden markov models in computational biology. applications to protein modeling. *J Mol Biol*, 235(5):1501–31, Feb 1994.

[29] Stan. Z. Li. *Markov Random Field Modeling in Computer Vision*. Springer-Verlag, 1995.

[30] Jinze Liu, Susan Paulsen, Wei Wang, Andrew Nobel, and Jan Prins. Mining approximate frequent itemsets from noisy data. In *Fifth IEEE International Conference on Data Mining (ICDM'05)*, pages 721–724, 2005.

[31] H. W. Mewes, C. Amid, R. Arnold, D. Frishman, U. Guldener, G. Mannhaupt, M. Munsterkotter, P. Pagel, N. Strack, V. Stumpflen, J. Warfsmann, and A. Ruepp. Mips: analysis and annotation of proteins from whole genomes. *Nucleic Acids Res*, 32 Database issue:D41–4, Jan 2004.

[32] H W Mewes, D Frishman, U Güldener, G Mannhaupt, K Mayer, M Mokrejs, B Morgenstern, M Münsterkötter, S Rudd, and B Weil. Mips: a database for genomes and protein sequences. *Nucleic Acids Res*, 30(1):31–34, Jan 2002. URL http://www.hubmed.org/display.cgi?uids=11752246.

[33] J Pei, A. K. Tung, and J. Han. Fault-tolerant frequent pattern mining: Problems and challenges. In *Workshop on Research Issues in Data Mining and Knowledge Discovery*, 2001.

[34] JB. Pereira-Leal, AJ. Enright, and CA. Ouzounis. Detection of functional modules from protein interaction networks. *Proteins*, 54(1):49–57, Jan 2004.

[35] R. Ramakrishnan and J. Gehrke. *Database Management Systems*. McGraw-Hill, 2006.

[36] T Reguly, A Breitkreutz, L Boucher, B J Breitkreutz, G C Hon, C L Myers, A Parsons, H Friesen, R Oughtred, A Tong, C Stark, Y Ho, D Botstein, B Andrews, C Boone, O G Troyanskya, T Ideker, K Dolinski, N N Batada, and M Tyers. Comprehensive curation and analysis of global interaction networks in saccharomyces cerevisiae. *J Biol*, 5(4):11–11, Jun 2006. URL `http://www.hubmed.org/display.cgi?uids=16762047`.

[37] D. Scholtens, M. Vidal, and R. Gentleman. Local modeling of global interactome networks. *Bioinformatics*, 21(17):3548–3557, Sep 2005.

[38] B Schwikowski, P Uetz, and S Fields. A network of protein-protein interactions in yeast. *Nat Biotechnol*, 18(12):1257–1261, Dec 2000.

[39] R. Sedgewick. *Algorithms in C*. Addison-Wesley, 1990.

[40] E Segal, H Wang, and D Koller. Discovering molecular pathways from protein interaction and gene expression data. *Bioinformatics*, 19 Suppl 1:264–271, 2003.

[41] B. Snel, P. Bork, and MA. Huynen. The identification of functional modules from the genomic association of genes. *Proc Natl Acad Sci U S A*, 99(9):5890–5895, Apr 2002.

[42] V. Spirin and LA. Mirny. Protein complexes and functional modules in molecular networks. *Proc Natl Acad Sci U S A*, 100(21):12123–12128, Oct 2003.

[43] X Sun and AB Nobel. Significance and recovery of block structure in binary matrices with noise. Technical report, Department of Statistics and Operation Research, UNC Chapel Hill, 2005.

[44] P Uetz. Two-hybrid arrays. *Curr Opin Chem Biol*, 6(1):57–62, Feb 2002. URL `http://www.hubmed.org/display.cgi?uids=11827824`.

[45] P Uetz, L Giot, G Cagney, T A Mansfield, R S Judson, J R Knight, D Lockshon, V Narayan, M Srinivasan, P Pochart, A Qureshi-Emili, Y Li, B Godwin, D Conover, T Kalbfleisch, G Vijayadamodar, M Yang, M Johnston, S Fields, and J M Rothberg. A comprehensive analysis of protein-protein interactions in saccharomyces cerevisiae. *Nature*, 403(6770):623–627, Feb 2000. URL `http://www.hubmed.org/display.cgi?uids=10688190`.

[46] Stijn van Dongen. *Graph Clustering by Flow Simulation*. PhD thesis, University of Utrecht, 2000.

[47] C. von Mering, R. Krause, B. Snel, M. Cornell, S. G. Oliver, S. Fields, and P. Bork. Comparative assessment of large-scale data sets of protein-protein interactions. *Nature*, 417(6887): 399–403, May 2002.

[48] Yongge Wang and Xintao Wu. Approximate inverse frequent itemset mining: Privacy, complexity, and approximation. In *Fifth IEEE International Conference on Data Mining (ICDM'05)*, pages 482–489, 2005.

[49] Cheng Yang, Usama Fayyad, and Paul S. Bradley. Efficient discovery of error-tolerant frequent itemsets in high dimensions. In *KDD '01: Proceedings of the seventh ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 194–203, New York, NY, USA, 2001. ACM Press. ISBN 1-58113-391-X.

[50] H Yu, A Paccanaro, V Trifonov, and M Gerstein. Predicting interactions in protein networks by completing defective cliques. *Bioinformatics*, 22(7):823–829, Apr 2006. URL `http://www.hubmed.org/display.cgi?uids=16455753`.