# References

F. Abramovich, Y. Benjamini, D. Donoho, and I. Johnstone. Adapting to unknown sparsity by controlling the false discovery rate. Technical Report 2000–19, Department of Statistics, Stanford University, Stanford, CA 94305, 2000.

F. Al-Shahrour, R. Díaz-Uriarte, and J. Dopazo. FatiGO: A web tool for finding significant associations of Gene Ontology terms with groups of genes. *Bioinformatics*, 20(4):578–580, 2004. Available at `fatigo.bioinfo.cipf.es`.

F. Al-Shahrour, P. Minguez, J. M. Vaquerizas, L. Conde, and J. Dopazo. BABELOMICS: A suite of web tools for functional annotation and analysis of groups of genes in high-throughput experiments. *Nucleic Acids Research*, 33:W460–W464, 2005. Available at `www.babelomics.org`.

A. A. Alizadeh, M. B. Eisen, R. E. Davis, C. Ma, I. S. Lossos, A. Rosenwald, J. C. Boldrick, H. Sabet, T. Tran, X. Yu, J. I. Powell, L. Yang, G. E. Marti, T. Moore, J. Hudson Jr, L. Lu, D. B. Lewis, R. Tibshirani, G. Sherlock, W. C. Chan, T. C. Greiner, D. D. Weisenburger, J. O. Armitage, R. Warnke, R. Levy, W. Wilson, M. R. Grever, J. C. Byrd, D. Botstein, P. O. Brown, and L. M. Staudt. Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. *Nature*, 403(6769):503–511, 2000.

B. Banelli, I. Casciano, M. Croce, A. Di Vinci, I. Gelvi, G. Pagnan, C. Brignole, G. Allemanni, S. Ferrini, M. Ponzoni, and M. Romani. Expression and methylation of CASP8 in neuroblastoma: Identification of a promoter region. *Nature Medicine*, 8(12):1333–1335, 2002.

J. D. Barbour, T. Wrin, R. M. Grant, J. N. Martin, M. R. Segal, C. J. Petropoulos, and S. G. Deeks. Evolution of phenotypic drug susceptibility and viral replication capacity during long-term virologic failure of protease inhibitor therapy in human immunodeficiency virus-infected adults. *Journal of Virology*, 76(21):11104–11112, 2002.

A. Barrier, P.-Y. Boëlle, A. Lemoine, C. Tse, D. Brault, F. Chiappini, F. Lacaine, S. Houry, M. Huguier, A. Flahault, and S. Dudoit. Gene expression profiling of nonneoplastic mucosa may predict clinical outcome of colon cancer patients. *Diseases of the Colon and Rectum*, 48(12):2238–2248, 2005a.

A. Barrier, A. Lemoine, P.-Y. Boëlle, C. Tse, D. Brault, F. Chiappini, J. Breittschneider, F. Lacaine, S. Houry, M. Huguier, M. J. van der Laan, T. P. Speed, B. Debuire, A. Flahault, and S. Dudoit. Colon cancer prognosis prediction by gene expression profiling. *Oncogene*, 24(40):6155–6164, 2005b.

A. Barrier, N. Olaya, F. Chiappini, F. Roser, O. Scatton, C. Artus, B. Franc, S. Dudoit, A. Flahault, B. Debuire, D. Azoulay, and A. Lemoine A. Ischemic preconditioning modulates the expression of several genes, leading to the overproduction of IL-1Ra, iNOS, and Bcl-2 in a human model of liver ischemia-reperfusion. *The FASEB Journal*, 19(12):1617–1626, 2005c.

A. Barrier, P.-Y. Boëlle, F. Roser, J. Gregg, C. Tse, D. Brault, F. Lacaine, S. Houry, M. Huguier, B. Franc, A. Flahault, A. Lemoine, and S. Dudoit. Stage II colon cancer prognosis prediction by tumor gene expression profiling. *Journal of Clinical Oncology*, 24(29):4685–4691, 2006.

T. Beissbarth and T. P. Speed. GOstat: Find statistically overrepresented Gene Ontologies within a group of genes. *Bioinformatics*, 20(9):1464–1465, 2004. Available at `gostat.wehi.edu.au`.

Y. Benjamini and H. Braun. John W. Tukey's contributions to multiple comparisons. *Annals of Statistics*, 30(6):1576–1594, 2002.

Y. Benjamini and Y. Hochberg. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society, Series B*, 57:289–300, 1995.

Y. Benjamini and D. Yekutieli. The control of the false discovery rate in multiple testing under dependency. *Annals of Statistics*, 29(4):1165–1188, 2001.

Y. Benjamini and D. Yekutieli. Quantitative trait loci analysis using the false discovery rate. *Genetics*, 171(2):783–790, 2005.

R. Beran. Balanced simultaneous confidence sets. *Journal of the American Statistical Association*, 83:679–686, 1988.

P. Billingsley. *Probability and Measure*. Probability and Mathematical Statistics. Wiley, New York, 2nd edition, 1986.

M. D. Birkner, A. E. Hubbard, and M. J. van der Laan. Application of a multiple testing procedure controlling the proportion of false positives to protein and bacterial data. Technical Report 186, Division of Biostatistics, University of California, Berkeley, Berkeley, CA 94720-7360, 2005a. Available at `www.bepress.com/ucbbiostat/paper186`.

M. D. Birkner, K. S. Pollard, M. J. van der Laan, and S. Dudoit. Multiple testing procedures and applications to genomics. Technical Report 168, Division of Biostatistics, University of California, Berkeley, Berkeley, CA 94720-7360, 2005b. Available at `www.bepress.com/ucbbiostat/paper168`.

M. D. Birkner, S. E. Sinisi, and M. J. van der Laan. Multiple testing and data adaptive regression: An application to HIV-1 sequence data. *Statistical Applications in Genetics and Molecular Biology*, 4(1):Article 8, 2005c. Available at `www.bepress.com/sagmb/vol4/iss1/art8`.

M. D. Birkner, A. E. Hubbard, M. J. van der Laan, C. F. Skibola, C. M. Hegedus, and M. T. Smith. Issues of processing and multiple testing of SELDI-TOF MS proteomic data. *Statistical Applications in Genetics and Molecular Biology*, 5(1):Article 11, 2006. Available at `www.bepress.com/sagmb/vol5/iss1/art11`.

M. D. Birkner, M. Courtine, M. J. van der Laan, K. Clément, J.-D. Zucker, and S. Dudoit. Statistical methods for detecting genotype/phenotype associations in the ObeLinks Project. Technical report, Division of Biostatistics, University of California, Berkeley, Berkeley, CA 94720-7360, 2007. (In preparation).

G. Bisson. KBG: A knowledge based generalizer. In B. W. Porter and R. J. Mooney, editors, *Machine Learning: Proceedings of the Seventh International Conference (1990), Austin, Texas, June 21–23, 1990*, pages 9–15, Palo Alto, CA, 1990. Morgan Kaufmann.

M. Blanchette, R. E. Green, S. E. Brenner, and D. C. Rio. Global analysis of positive and negative pre-mRNA splicing regulators in *Drosophila*. *Genes & Development*, 19(11):1306–1314, 2005.

J. C. Boldrick, A. A. Alizadeh, M. Diehn, S. Dudoit, C. L. Liu, C. E. Belcher, D. Botstein, L. M. Staudt, P. O. Brown, and D. A. Relman. Stereotyped and specific gene expression programs in human innate immune responses to bacteria. *Proc. Natl. Acad. Sci.*, 99(2):972–977, 2002.

B. M. Bolstad, R. A. Irizarry, L. Gautier, and Z. Wu. Preprocessing high-density oligonucleotide arrays. In R. C. Gentleman, V. J. Carey, W. Huber, R. A. Irizarry, and S. Dudoit, editors, *Bioinformatics and Computational Biology Solutions Using R and Bioconductor*, chapter 2, pages 13–32. Springer, New York, 2005. Available at `www.bioconductor.org/pubs/docs/mogr`.

C. E. Bonferroni. Teoria statistica delle classi e calcolo delle probabilità. *Pubblicazioni del R Istituto Superiore di Scienze Economiche e Commerciali di Firenze*, pages 3–62, 1936.

J. P. Bordat. Calcul pratique du treillis de Galois d'une correspondance. *Mathématique des Sciences Humaines*, 96(4):31–47, 1986.

I. Bournaud, M. Courtine, and J.-D. Zucker. KIDS: An iterative algorithm to organize relational knowledge. In R. Dieng and O. Corby, editors, *Knowledge Engineering and Knowledge Management. Methods, Models, and Tools: 12th International Conference, EKAW 2000, Juan-les-Pins, France, October 2–6, 2000, Proceedings*, volume 1937 of *Lecture Notes in Computer Science*, pages 217–232, Berlin/Heidelberg, 2000. Springer. Available at `www.springerlink.com/content/472a404urcgxjqn9/?p=db5cfd25a3c54db7a2dcf7c2e1cfd3fc&pi=15`.

G. A. Calin, C. Sevignani, C. D. Dumitru, T. Hyslop, E. Noch, S. Yendamuri, M. Shimizu, S. Rattan, F. Bullrich, M. Negrini, and C. M. Croce. Human microRNA genes are frequently located at fragile sites and genomic regions involved in cancers. *Proc. Natl. Acad. Sci.*, 101(9):2999–3004, 2004.

M. J. Callow, S. Dudoit, E. L. Gong, T. P. Speed, and E. M. Rubin. Microarray expression profiling identifies genes with altered expression in HDL deficient mice. *Genome Research*, 10(12):2022–2029, 2000.

C. Carpineto and G. Romano. A lattice conceptual clustering system and its application to browsing retrieval. *Machine Learning*, 24(2):95–122, 1996.

S. Cawley, S. Bekiranov, H. H. Ng, P. Kapranov, E. A. Sekinger, D. Kampa, A. Piccolboni, V. Sementchenko, J. Cheng, A. J. Williams, R. Wheeler, B. Wong, J. Drenkow, M. Yamanaka, S. Patel, S. Brubaker, H. Tammana, G. Helt, K. Struhl, and T. R. Gingeras. Unbiased mapping of transcription factor binding sites along human chromosomes 21 and 22 points to widespread regulation of noncoding RNAs. *Cell*, 116:499–509, 2004.

J. M. Chambers. *Programming with Data.* Springer, New York, 1998.

M. Chein. Algorithmes de recherche des sous-matrices premières d'une matrice. *Bull. Math. Soc. Sci. Math. RS Roumanie*, 13(61):21–25, 1969.

A. M. Cheng, M. W. Byrom, J. Shelton, and L. P. Ford. Antisense inhibition of human miRNAs and indications for an involvement of miRNA in cell growth and apoptosis. *Nucleic Acids Research*, 33(4):1290–1297, 2005.

F. Chiappini, A. Barrier, R. Saffroy, M.-C. Domart, N. Dagues, D. Azoulay, M. Sebagh, B. Franc, S. Chevalier, B. Debuire, S. Dudoit, and A. Lemoine. Exploration of global gene expression in human liver steatosis by high-density oligonucleotide microarray. *Laboratory Investigation*, 86(2):154–165, 2006.

S. Chiaretti, X. Li, R. Gentleman, A. Vitale, M. Vignetti, F. Mandelli, J. Ritz, and R. Foa. Gene expression profile of adult T-cell acute lymphocytic leukemia identifies distinct subsets of patients with different response to therapy and survival. *Blood*, 103(7):2771–2778, 2004.

K. Clément. Genetics of human obesity. *Proc. Nutr. Soc.*, 64(2):133–142, 2005.

K. Clément and P. Ferré. Genetics and the pathophysiology of obesity. *Pediatric Research*, 53(5):721–725, 2003.

M. J. Daly, J. D. Rioux, S. F. Schaffner, T. J. Hudson, and E. S. Lander. High-resolution haplotype structure in the human genome. *Nature Genetics*, 29 (2):229–232, 2001.

T. Z. DeSantis, C. E. Stone, S. R. Murray, J. P. Moberg, and G. L. Andersen. Rapid quantification and taxonomic classification of environmental DNA from both prokaryotic and eukaryotic origins using a microarray. *FEMS Microbiology Letters*, 245(2):271–278, 2005.

S. Dudoit and Y. H. Yang. Bioconductor R packages for exploratory analysis and normalization of cDNA microarray data. In G. Parmigiani, E. S. Garrett, R. A. Irizarry, and S. L. Zeger, editors, *The Analysis of Gene Expression Data: Methods and Software*, pages 73–101. Springer, New York, 2003.

S. Dudoit, Y. H. Yang, M. J. Callow, and T. P. Speed. Statistical methods for identifying differentially expressed genes in replicated cDNA microarray experiments. *Statistica Sinica*, 12(1):111–139, 2002.

S. Dudoit, J. P. Shaffer, and J. C. Boldrick. Multiple hypothesis testing in microarray experiments. *Statistical Science*, 18(1):71–103, 2003.

S. Dudoit, M. J. van der Laan, and M. D. Birkner. Multiple testing procedures for controlling tail probability error rates. Technical Report 166, Division of Biostatistics, University of California, Berkeley, Berkeley, CA 94720-7360, 2004a. Available at `www.bepress.com/ucbbiostat/paper166`.

S. Dudoit, M. J. van der Laan, and K. S. Pollard. Multiple testing. Part I. Single-step procedures for control of general Type I error rates. *Statistical Applications in Genetics and Molecular Biology*, 3(1):Article 13, 2004b. Available at `www.bepress.com/sagmb/vol3/iss1/art13`.

S. Dudoit, S. Keleş, and M. J. van der Laan. Multiple tests of association with biological annotation metadata. Technical Report 202, Division of Biostatistics, University of California, Berkeley, Berkeley, CA 94720-7360, 2006. Available at `www.bepress.com/ucbbiostat/paper202`.

O. J. Dunn. Estimation of the means of dependent variables. *Annals of Mathematical Statistics*, 29:1095–111, 1958.

B. Efron. Local false discovery rates. Technical report, Department of Statistics, Stanford University, Stanford, CA 94305, 2005. Available at `www-stat.stanford.edu/~brad/papers`.

B. Efron, J. D. Storey, and R. Tibshirani. Microarrays, empirical Bayes methods, and false discovery rates. Technical Report 2001–218, Department of Statistics, Stanford University, Stanford, CA 94305, 2001a.

B. Efron, R. Tibshirani, J. D. Storey, and V. Tusher. Empirical Bayes analysis of a microarray experiment. *Journal of the American Statistical Association*, 96:1151–1160, 2001b.

H. Finner. Stepwise multiple test procedures and control of directional errors. *Annals of Statistics*, 27:274–289, 1999.

S. B. Gabriel, S. F. Schaffner, H. Nguyen, J. M. Moore, J. Roy, B. Blumenstiel, J. Higgins, M. DeFelice, A. Lochner, M. Faggart, S. N. Liu-Cordero, C. Rotimi, A. Adeyemo, R. Cooper, R. Ward, E. S. Lander, M. J. Daly, and D. Altshuler. The structure of haplotype blocks in the human genome. *Science*, 296:2225–2229, 2002.

B. Ganter. Two basic algorithms in concept analysis. Technical Report 831, Technische Hochschule, Darmstadt, Germany, 1984.

Y. Ge, S. Dudoit, and T. P. Speed. Resampling-based multiple testing for microarray data analysis. *TEST*, 12(1):1–44, 2003.

C. R. Genovese and L. Wasserman. A stochastic process approach to false discovery control. *Annals of Statistics*, 32(3):1035–1061, 2004a.

C. R. Genovese and L. Wasserman. Exceedance control of the false discovery proportion. Technical Report 807, Department of Statistics, Carnegie Mellon University, Pittsburgh, PA 15213, July 2004b. Available at `www.stat.cmu.edu/tr/tr807/tr807.html`.

C. R. Genovese, N. A. Lazar, and T. E. Nichols. Thresholding of statistical maps in functional neuroimaging using the false discovery rate. *NeuroImage*, 15(4):870–878, 2002.

R. C. Gentleman, V. J. Carey, D. J. Bates, B. Bolstad, M. Dettling, S. Dudoit, B. Ellis, L. Gautier, Y. Ge, J. Gentry, K. Hornik, T. Hothorn, W. Huber, S. Iacus, R. A. Irizarry, F. Leisch, C. Li, M. Maechler, A. J. Rossini, G. Sawitzki, C. Smith, G. K. Smyth, L. Tierney, Y. H. Yang, and J. Zhang. Bioconductor: Open software development for computational biology and bioinformatics. *Genome Biology*, 5(10):R80, 2004. Available at `genomebiology.com/2004/5/10/R80,www.bioconductor.org`.

R. C. Gentleman, V. J. Carey, W. Huber, R. A. Irizarry, and S. Dudoit, editors. *Bioinformatics and Computational Biology Solutions Using R and Bioconductor.* Statistics for Biology and Health. Springer, New York, 2005a. Available at `www.bioconductor.org/pubs/docs/mogr`.

R. C. Gentleman, V. J. Carey, and J. Zhang. Meta-data resources and tools in Bioconductor. In R. C. Gentleman, V. J. Carey, W. Huber, R. A. Irizarry, and S. Dudoit, editors, *Bioinformatics and Computational Biology Solutions Using R and Bioconductor*, chapter 7, pages 111–133. Springer, New York, 2005b. Available at `www.bioconductor.org/pubs/docs/mogr`.

R. D. Gill. Non- and semi-parametric maximum likelihood estimators and the von Mises method. I. *Scandinavian Journal of Statistics*, 16(2):97–128, 1989. (With a discussion by J. A. Wellner and J. Præstgaard and a reply by the author).

R. D. Gill, M. J. van der Laan, and J. A. Wellner. Inefficient estimators of the bivariate survival function for three models. *Annales de l'Institut Henri Poincaré. Probabilités et Statistiques*, 31(3):545–597, 1995.

B. Gleissner, N. Gokbuget, C. R. Bartram, B. Janssen, H. Rieder, J. W. Janssen, C. Fonatsch, A. Heyll, D. Voliotis, J. Beck, T. Lipp, G. Munzert, J. Maurer, D. Hoelzer, E. Thiel, and German Multicenter Trials of Adult Acute Lymphoblastic Leukemia Study Group. Leading prognostic relevance of the BCR-ABL translocation in adult acute B-lineage lymphoblastic leukemia: A prospective study of the German Multicenter Trial Group and confirmed polymerase chain reaction analysis. *Blood*, 99(5):1536–1543, 2002.

A. L. Gloyn, M. N. Weedon, K. R. Owen, M. J. Turnerand B. A. Knight, G. Hitman, M. Walker, J. C. Levy, M. Sampson, S. Halford, M. I. McCarthy, A. T. Hattersley, and T. M. Frayling. Large-scale association studies of variants in genes encoding the pancreatic $\beta$-cell $K_{ATP}$ channel subunits Kir6.2 (*KCNJ11*) and SUR1 (*ABCC8*) confirm that the *KCNJ11* E23K variant is associated with type 2 diabetes. *Diabetes*, 52(2):568–572, 2003.

R. Godin, G. Mineau, R. Missaoui, and H. Mili. Méthodes de classification conceptuelle basées sur les treillis de Galois et applications. *Revue d'Intelligence Artificielle*, 9(2):105–137, 1995a.

R. Godin, R. Missaoui, and H. Alaoui. Incremental concept formation algorithms based on Galois (concept) lattices. *Computational Intelligence*, 11 (2):216–267, 1995b.

T. R. Golub, D. K. Slonim, P. Tamayo, C. Huard, M. Gaasenbeek, J. P. Mesirov, H. Coller, M.L. Loh, J. R. Downing, M. A. Caligiuri, C. D. Bloomfield, and E. S. Lander. Molecular classification of cancer: Class discovery and class prediction by gene expression monitoring. *Science*, 286(5439): 531–537, 1999.

V. Goss Tusher, R. Tibshirani, and G. Chu. Significance analysis of microarrays applied to the ionizing radiation response. *Proc. Natl. Acad. Sci.*, 98: 5116–5121, 2001.

S. Grossmann, S. Bauer, P. N. Robinson, and M. Vingron. An improved statistic for detecting over-represented gene ontology annotations in gene sets. In A. Apostolico, C. Guerra, S. Istrail, P. Pevzner, and M. Waterman, editors, *Research in Computational Molecular Biology: 10th Annual International Conference, RECOMB 2006, Venice, Italy, April 2–5, 2006, Proceedings*, volume 3909 of *Lecture Notes in Computer Science*, pages 85–98, Berlin/Heidelberg, 2006. Springer. Available at `www.springerlink.com/content/w83h54235ku1v142/?p=a2cb65a9ec484e519ad1ff83b41c707a&pi=8`.

L. W. Hahn, M. D. Ritchie, and J. H. Moore. Multifactor dimensionality reduction software for detecting gene-gene and gene-environment interactions. *Bioinformatics*, 19(3):376–382, 2003.

E. H. Hani, K. Clément, G. Velho, N. Vionnet, J. Hager, A. Philippi, C. Dina, H. Inoue, M. A. Permutt, A. Basdevant, M. North, F. Demenais, B. Guy-Grand, and P. Froguel. Genetic studies of the sulfonylurea receptor gene locus in NIDDM and in morbid obesity among French Caucasians. *Diabetes*, 46(4):688–694, 1997.

E. R. Hauser and M. Boehnke. Genetic linkage analysis of complex genetic traits by using affected sibling pairs. *Biometrics*, 54:1238–1246, 1998.

A. Herbert, N. P. Gerry, M. B. McQueen, I. M. Heid, A. Pfeufer, T. Illig, H.-E. Wichmann, T. Meitinger, D. Hunter, F. B. Hu, G. Colditz, A. Hinney, J. Hebebrand, K. Koberwitz, X. Zhu, R. Cooper, K. Ardlie, H. Lyon, J. N. Hirschhorn, N. M. Laird, M. E. Lenburg, C. Lange, and M. F. Christman. A common genetic variant is associated with adult and childhood obesity. *Science*, 312(5771):279–283, 2006.

A. von Heydebreck, W. Huber, and R. Gentleman. Differential expression with the Bioconductor Project. Technical Report 7, Bioconductor Project Working Papers, 2004. Available at `www.bepress.com/bioconductor/paper7`.

Y. Hochberg. A sharper Bonferroni procedure for multiple tests of significance. *Biometrika*, 75:800–802, 1988.

Y. Hochberg and A. C. Tamhane. *Multiple Comparison Procedures*. Probability and Mathematical Statistics. Wiley, New York, 1987.

R. P. Hoffman, P. Vicini, and C. Cobelli. Pubertal changes in HOMA and QUICKI: Relationship to hepatic and peripheral insulin sensitivity. *Pediatric Diabetes*, 5(3):122–125, 2004.

J. Hoh and J. Ott. Mathematical multi-locus approaches to localizing complex human trait genes. *Nature Reviews Genetics*, 4(9):701–709, 2003.

J. Hoh and J. Ott. Genetic dissection of diseases: Design and methods. *Current Opinion in Genetics & Development*, 14(3):229–232, 2004.

B. Holland and M. D. Copenhaver. An improved sequentially rejective Bonferroni test procedure. *Biometrics*, 43(2):417–423, 1987.

S. Holm. A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, 6:65–70, 1979.

G. Hommel. Tests of the overall hypothesis for arbitrary dependence structures. *Biometrical Journal*, 25:423–430, 1983.

G. Hommel. A stagewise rejective multiple test procedure based on a modified Bonferroni test. *Biometrika*, 75:383–386, 1988.

G. Hommel and T. Hoffman. Controlled uncertainty. In P. Bauer, G. Hommel, and E. Sonnemann, editors, *Multiple Hypotheses Testing*, pages 154–161. Springer, 1988.

J. C. Hsu. *Multiple Comparisons: Theory and Methods*. Chapman & Hall/CRC, London, New York, 1996. Available at `www.stat.ohio-state.edu/~jch`.

Y. Huang and J. C. Hsu. Hochberg's step-up method: Cutting corners off Holm's step-down method. Technical Report 761, Department of Statistics, Ohio State University, Columbus, OH 43210-1247, 2005.

K. Jogdeo. Association and probability inequalities. *Annals of Statistics*, 5: 495–504, 1977.

J. Kaczynski, T. Cook, and R. Urrutia. Sp1- and Krüppel-like transcription factors. *Genome Biology*, 4(2):206, 2003.

T. Kamae, U. Krengel, and G. L. O'Brien. Stochastic inequalities on partially ordered spaces. *Annals of Probability*, 5(6):899–912, 1977.

S. Keleş, M. J. van der Laan, S. Dudoit, and S. E. Cawley. Multiple testing methods for ChIP-Chip high density oligonucleotide array data. *Journal of Computational Biology*, 13(3):579–613, 2006. Available at `www.liebertonline.com/doi/abs/10.1089/cmb.2006.13.579?prevSearch=allfield%3A%28dudoit%29`.

D. Kirchner, J. Duyster, O. Ottmann, R. M. Schmid, L. Bergmann, and G. Munzert. Mechanisms of Bcr-Abl-mediated NF-$\kappa$B/Rel activation. *Experimental Hematology*, 31(6):504–511, 2003.

E. L. Korn, J. F. Troendle, L. M. McShane, and R. Simon. Controlling the number of false discoveries: Application to high-dimensional genomic data. *Journal of Statistical Planning and Inference*, 124(2):379–398, 2004.

S. O. Kuznetsov and S. A. Obiedkov. Comparing performance of algorithms for generating concept lattices. *Journal of Experimental and Theoretical Artificial Intelligence*, 14(2–3):189–216, 2002.

M. J. van der Laan. Statistical inference for variable importance. *International Journal of Biostatistics*, 2(1):Article 2, 2006. Available at `www.bepress.com/ijb/vol2/iss1/2`.

M. J. van der Laan and A. E. Hubbard. Quantile-function based null distribution in resampling based multiple testing. *Statistical Applications in Genetics and Molecular Biology*, 5(1):Article 14, 2006. Available at `www.bepress.com/sagmb/vol5/iss1/art14`.

M. J. van der Laan and K. S. Pollard. A new algorithm for hybrid hierarchical clustering with visualization and the bootstrap. *Journal of Statistical Planning and Inference*, 117(2):275–303, 2003.

M. J. van der Laan and J. M. Robins. *Unified Methods for Censored Longitudinal Data and Causality.* Springer, New York, 2003.

M. J. van der Laan, S. Dudoit, and K. S. Pollard. Multiple testing. Part II. Step-down procedures for control of the family-wise error rate. *Statistical Applications in Genetics and Molecular Biology*, 3(1):Article 14, 2004a. Available at `www.bepress.com/sagmb/vol3/iss1/art14`.

M. J. van der Laan, S. Dudoit, and K. S. Pollard. Augmentation procedures for control of the generalized family-wise error rate and tail probabilities for the proportion of false positives. *Statistical Applications in Genetics and Molecular Biology*, 3(1):Article 15, 2004b. Available at `www.bepress.com/sagmb/vol3/iss1/art15`.

M. J. van der Laan, M. D. Birkner, and A. E. Hubbard. Empirical Bayes and resampling based multiple testing procedure controlling tail probability of the proportion of false positives. *Statistical Applications in Genetics and Molecular Biology*, 4(1):Article 29, 2005. Available at `www.bepress.com/sagmb/vol4/iss1/art29`.

E. L. Lehmann. *Testing Statistical Hypotheses.* Wiley, New York, 2nd edition, 1986.

E. L. Lehmann and J. P. Romano. Generalizations of the familywise error rate. *Annals of Statistics*, 33(3):1138–1154, 2005.

S. S. Li, J. Bigler, J. W. Lampe, J. D. Potter, and Z. Feng. FDR-controlling testing procedures and sample size determination for microarrays. *Statistics in Medicine*, 24(15):2267–2280, 2005.

J. Lu, G. Getz, E. A. Miska, E. Alvarez-Saavedra, J. Lamb, D. Peck, A. Sweet-Cordero, B. L. Ebert, R. H. Mak, A. A. Ferrando, J. R. Downing, T. jacks, H. R. Horvitz, and T. R. Golub. MicroRNA expression profiles classify human cancers. *Nature*, 435(9):834–838, 2005. Available at `www.broad.mit.edu/cancer/pub/miGCM`.

D. Maglott, J. Ostell, K. D. Pruitt, and T. Tatusova. Entrez Gene: Gene-centered information at NCBI. *Nucleic Acids Research*, 33:D54–D58, 2005. Available at `nar.oxfordjournals.org/cgi/content/full/33/suppl_1/D54`.

E. Manduchi, G. R. Grant, S. E. McKenzie, G. C. Overton, S. Surrey, and C. J. Stoeckert. Generation of patterns from gene expression data by assigning confidence to differentially expressed genes. *Bioinformatics*, 16:685–698, 2000.

K. V. Mardia, J. T. Kent, and J. M. Bibby. *Multivariate Analysis.* Academic Press, London, New York, 1979.

S. A. McCarroll, C. T. Murphy, S. Zou, S. D. Pletcher, C-S Chin, Y. N. Jan, C. Kenyon, C. I. Bargmann, and H. Li. Comparing genomic expression patterns across species identifies shared transcriptional profile in aging. *Nature Genetics*, 36(2):197–204, 2004.

E. Mephu Nguifo and P. Njiwoua. Using lattice-based framework as a tool for feature extraction. In C. Nédellec and C. Rouveirol, editors, *Machine Learning: ECML-98: 10th European Conference on Machine Learning, Chemnitz, Germany, April 21–23, 1998, Proceedings*, volume 1398 of *Lecture Notes in Computer Science*, pages 304–309, Berlin/Heidelberg, 1998. Springer.

D. Meyre, N. Bouatia-Naji, A. Tounian, C. Samson, C. Lecoeur, V. Vatin, M. Ghoussaini, C. Wachter, S. Hercberg, G. Charpentier, W. Patsch, F. Pattou, M.-A. Charles, P. Tounian, K. Clément, B. Jouret, J. Weill, B. A. Maddux, I. D. Goldfine, A. Walley, P. Boutin, C. Dina, and P. Froguel. Variants of ENPP1 are associated with childhood and adult obesity and increase the risk of glucose intolerance and type 2 diabetes. *Nature Genetics*, 37(8): 863–867, 2005. Available at `www.nature.com/ng/journal/v37/n8/abs/ng1604.html;jsessionid=741AF577246F06B4483ED78D3CAD091A`.

M. Z. Michael, S. M. O'Connor, N. G. van Holst Pellekaan, G. P. Young, and R. J. James. Reduced accumulation of specific microRNAs in colorectal neoplasia. *Molecular Cancer Research*, 1(12):882–891, 2003.

V. K. Mootha, C. M. Lindgren, K-F. Eriksson, A. Subramanian, S. Sihag, J. Lehar, P. Puigserver, E. Carlsson, M. Ridderstråle, E. Laurila, N. Houstis, M. J. Daly, N. Patterson, J. P. Mesirov, T. R. Golub, P. Tamayo, B. Spiegelman, E. S. Lander, J. H. Hirschhorn, D. Altshuler, and L. C. Groop. PGC–1$\alpha$-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. *Nature Genetics*, 34(3):267–273, 2003.

A. Mukhopadhyay, S. Shishodia, J. Suttles, K. Brittingham, B. Lamothe, R. Nimmanapalli, K. N. Bhalla, and B. B. Aggarwal. Ectopic expression of protein-tyrosine kinase Bcr-Abl suppresses tumor necrosis factor (TNF)-induced NF-$\kappa$B activation and I$\kappa$B$\alpha$ phosphorylation. Relationship with down-regulation of TNF receptors. *Journal of Biological Chemistry*, 277 (34):30622–30628, 2002.

M. A. Newton, C. M. Kendziorski, C. S. Richmond, F. R . Blattner, and K. W. Tsui. On differential variability of expression ratios: Improving statistical inference about gene expression changes from microarray data. *Journal of Computational Biology*, 8:37–52, 2001.

E. M. Norris. An algorithm for computing the maximal rectangles in a binary relation. *Revue Roumaine de Mathématiques Pures et Appliquées*, 23(2): 243–250, 1978.

K. A. O'Donnell, E. A. Wentzel, K. I. Zeller, C. V. Dang, and J. T. Mendell. c-Myc-regulated microRNAs modulate E2F1 expression. *Nature*, 435(7043): 839–843, 2005.

J. M. Ordovas. Genetics, postprandial lipemia and obesity. *Nutrition, Metabolism & Cardiovascular Diseases*, 11(2):118–133, 2001.

A. Packer, editor. *The Chipping Forecast II*, volume 32(4s) of *Nature Genetics*. Nature Publishing Group, December 2002. Available at `www.nature.com/ng/journal/v32/n4s`. (Supplement).

A. Packer and M. Axton, editors. *The Chipping Forecast III*, volume 37(6s) of *Nature Genetics*. Nature Publishing Group, June 2005. Available at `www.nature.com/ng/journal/v37/n6s`. (Supplement).

L. Pérusse, T. Rankinen, A. Zuberi, Y. C. Chagnon, S. J. Weisnagel, G. Argyropoulos, B. Walts, E. E. Snyder, and C. Bouchard. The human obesity gene map: The 2004 update. *Obesity Research*, 13(3):381–490, 2005.

E. Phimister and B. Cohen, editors. *The Chipping Forecast*, volume 21(1s) of *Nature Genetics*. Nature of America, January 1999. Available at `www.nature.com/ng/journal/v21/n1s`. (Supplement).

K. S. Pollard and M. J. van der Laan. Choice of a null distribution in resampling-based multiple testing. *Journal of Statistical Planning and Inference*, 125(1–2):85–100, 2004.

K. S. Pollard and M. J. van der Laan. Cluster analysis of genomic data. In R. C. Gentleman, V. J. Carey, W. Huber, R. A. Irizarry, and S. Dudoit, editors, *Bioinformatics and Computational Biology Solutions Using R and Bioconductor*, chapter 13, pages 209–228. Springer, New York, 2005. Available at `www.bioconductor.org/pubs/docs/mogr,www.bepress.com/ucbbiostat/paper167`.

K. S. Pollard, M. D. Birkner, M. J. van der Laan, and S. Dudoit. Test statistics null distributions in multiple testing: Simulation studies and applications to genomics. *Journal de la Société Française de Statistique*, 146(1–2):77–115, 2005a. Available at `www.stat.berkeley.edu/~sandrine/Docs/Papers/SFdS05/SFdS.html`. Numéro double spécial *Statistique et Biopuces.*

K. S. Pollard, S. Dudoit, and M. J. van der Laan. Multiple testing procedures: The multtest package and applications to genomics. In R. C. Gentleman, V. J. Carey, W. Huber, R. A. Irizarry, and S. Dudoit, editors, *Bioinformatics and Computational Biology Solutions Using R and Bioconductor*, chapter 15, pages 249–271. Springer, New York, 2005b. Available at `www.bioconductor.org/pubs/docs/mogr,www.bepress.com/ucbbiostat/paper164`.

M. Prochazka, S. Lillioja, J. F. Tait, W. C. Knowler, D. M. Mott, M. Spraul, P. H. Bennett, and C. Bogardus. Linkage of chromosomal markers on 4q with a putative gene determining maximal insulin action in Pima Indians. *Diabetes*, 42(4):514–519, 1993.

M. L. Puri and P. K. Sen. *Nonparametric Methods in Multivariate Analysis.* Wiley, New York, 1971.

R Development Core Team. *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing, Vienna, Austria, 2006. Available at `www.r-project.org`.

P. H. Ramsey. Power differences between pairwise multiple comparisons. *Journal of the American Statistical Association*, 73:479–485, 1978.

A. Reiner, D. Yekutieli, and Y. Benjamini. Identifying differentially expressed genes using false discovery rate controlling procedures. *Bioinformatics*, 19 (3):368–375, 2003.

M. F. Rolland-Cachera, T. J. Cole, M. Sempe, J. Tichet, C. Rossignol, and A. Charraud. Body Mass Index variations: Centiles from birth to 87 years. *European Journal of Clinical Nutrition*, 45(1):13–21, 1991.

D. M. Rom. A sequentially rejective test procedure based on a modified Bonferroni inequality. *Biometrika*, 77:663–665, 1990.

J. P. Romano and M. Wolf. Control of generalized error rates in multiple testing. Technical Report 2005–12, Department of Statistics, Stanford University, Stanford, CA 94305, 2005.

D. Rubin, M. J. van der Laan, and S. Dudoit. A method to increase the power of multiple testing procedures through sample splitting. *Statistical Applications in Genetics and Molecular Biology*, 5(1):Article 19, 2006. Available at `www.bepress.com/sagmb/vol5/iss1/art19`.

S. K. Sarkar. Some results on false discovery rate in stepwise multiple testing procedures. *Annals of Statistics*, 30(1):239–257, 2002.

S. K. Sarkar. Generalizing Simes' test and Hochberg's stepup procedure. Technical report, Fox School of Business and Management, Temple University, Philadelphia, PA 19122, August 2005.

S. K. Sarkar. Probability inequalities for ordered MTP2 random variables: A proof of the Simes conjecture. *Annals of Statistics*, 26(2):494–504, 1998.

S. K. Sarkar and C-K. Chang. The Simes method for multiple hypothesis testing with positively dependent test statistics. *Journal of the American Statistical Association*, 92:1601–1608, 1997.

S. R. Searle. *Linear Models*. Wiley, New York, 1971.

P. Seeger. A note on a method for the analysis of significances en masse. *Technometrics*, 10(3):586–593, 1968.

M. R. Segal, J. D. Barbour, and R. M. Grant. Relating HIV-1 sequence variation to replication capacity via trees and forests. *Statistical Applications in Genetics and Molecular Biology*, 3(1):Article 2, 2004. Available at `www.bepress.com/sagmb/vol3/iss1/art2`.

R. W. Shafer, K. M. Dupnik, M. A. Winters, and S. H. Eshleman. A guide to HIV-1 reverse transcriptase and protease sequencing for drug resistance studies. In *HIV Sequencing Compendium*, pages 83–133. Theoretical Biology and Biophysics Group, Los Alamos National Laboratory, 2001.

J. P. Shaffer. Multiplicity, directional (Type III) errors, and the null hypothesis. *Psychological Methods*, 7:356–369, 2002.

J. P. Shaffer. Modified sequentially rejective multiple test procedures. *Journal of the American Statistical Association*, 81:826–831, 1986.

J. P. Shaffer. Multiple hypothesis testing. *Annu. Rev. Psychol.*, 46:561–584, 1995.

E. Shtivelman, F. E. Cohen, and J. M. Bishop. A human gene (AHNAK) encoding an unusually large protein with a 1.2-$\mu$m polyionic rod structure. *Proc. Natl. Acad. Sci.*, 89(12):5472–5476, 1992.

Z. Šidák. Rectangular confidence regions for the means of multivariate normal distributions. *Journal of the American Statistical Association*, 62:626–633, 1967.

R. J. Simes. An improved Bonferroni procedure for multiple tests of significance. *Biometrika*, 73:751–754, 1986.

S. E. Sinisi and M. J. van der Laan. Deletion/substitution/addition algorithm in learning with applications in genomics. *Statistical Applications in Genetics and Molecular Biology*, 3(1):Article 18, 2004. Available at `www.bepress.com/sagmb/vol3/iss1/art18`.

B. Sorić. Statistical "discoveries" and effect-size estimation. *Journal of the American Statistical Association*, 84(406):608–610, 1989.

T. P. Speed, editor. *Statistical Analysis of Gene Expression Microarray Data.* Chapman & Hall/CRC, Boca Raton, FL, 2003.

J. D. Storey. The positive false discovery rate: A Bayesian interpretation and the q-value. *Annals of Statistics*, 31(6):2013–2035, 2003.

J. D. Storey. A direct approach to false discovery rates. *Journal of the Royal Statistical Society, Series B*, 64(3):479–498, 2002.

J. D. Storey and R. Tibshirani. SAM thresholding and false discovery rates for detecting differential gene expression in DNA microarrays. In G. Parmigiani, E. S. Garrett, R. A. Irizarry, and S. L. Zeger, editors, *The Analysis of Gene Expression Data: Methods and Software*, pages 272–290. Springer, New York, 2003.

J. D. Storey and R. Tibshirani. Estimating false discovery rates under dependence, with applications to DNA microarrays. Technical Report 2001–28, Department of Statistics, Stanford University, Stanford, CA 94305, 2001.

J. D. Storey, J. E. Taylor, and D. O. Siegmund. Strong control, conservative point estimation and simultaneous conservative consistency of false discovery rates: A unified approach. *Journal of the Royal Statistical Society, Series B*, 66(1):187–205, 2004.

A. Subramanian, P. Tamayo, V. K. Mootha, S. Mukherjee, B. L. Ebert, M. A. Gillette, A. Paulovich, S. L. Pomeroy, T. R. Golub, E. S. Lander, and J. P. Mesirov. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci.*, 102 (43):15545–15550, 2005. Available at `www.broad.mit.edu/gsea/doc/doc_index.html`.

M. M. Swarbrick and C. Vaisse. Emerging trends in the search for genetic variants predisposing to human obesity. *Current Opinion in Clinical Nutrition and Metabolic Care*, 6(4):369–375, 2003.

L. Tian, S. A. Greenberg, S. W. Kong, J. Altschuler, I. S. Kohane, and P. J. Park. Discovering statistically significant pathways in expression profiling studies. *Proc. Natl. Acad. Sci.*, 102(38):13544–13549, 2005.

J. F. Troendle. A stepwise resampling method of multiple hypothesis testing. *Journal of the American Statistical Association*, 90(429):370–378, 1995.

J. F. Troendle. A permutational step-up method of testing multiple outcomes. *Biometrics*, 52:846–859, 1996.

M. Tsunoda, J. Tenhunen, C. Tilgmann, H. Arai, and K. Imai. Reduced membrane-bound catechol-O-methyltransferase in the liver of spontaneously hypertensive rats. *Hypertension Research*, 26(11):923–927, 2003.

A. W. van der Vaart and J. Wellner. *Weak Convergence and Empirical Processes*. Springer, New York, 1996.

Y. Wang and S. Dudoit. Quantification and visualization of LD patterns and identification of haplotype blocks. Technical Report 150, Division of Biostatistics, University of California, Berkeley, Berkeley, CA 94720-7360, 2004. Available at `www.bepress.com/ucbbiostat/paper150`.

P. H. Westfall. Resampling-based multiple testing for microarray data analysis. *TEST*, 12(1), 2003. (Discussion).

P. H. Westfall and S. S. Young. *Resampling-Based Multiple Testing: Examples and Methods for P-Value Adjustment*. Wiley, New York, 1993.

E. Wienholds and R. H. A. Plasterk. MicroRNA function in animal development. *Federation of European Biochemical Societies Letters*, 579(26): 5911–5922, 2005.

R. Wille. Restructuring lattice theory: An approach based on hierarchies of concepts. In I. Rival, editor, *Ordered Sets*, volume 83, pages 445–470. D. Reidel, Dordrecht-Boston, 1982.

R. Wille. Concept lattices and conceptual knowledge systems. *Computers & Mathematics with Application*, 23(6–9):493–515, 1992.

Y. H. Yang and A. C. Paquet. Preprocessing two-color spotted arrays. In R. C. Gentleman, V. J. Carey, W. Huber, R. A. Irizarry, and S. Dudoit, editors, *Bioinformatics and Computational Biology Solutions Using R and Bioconductor*, chapter 4, pages 49–69. Springer, New York, 2005. Available at `www.bioconductor.org/pubs/docs/mogr`.

Y. H. Yang, S. Dudoit, P. Luu, and T. P. Speed. Normalization for cDNA microarray data. In M. L. Bittner, Y. Chen, A. N. Dorsel, and E. R. Dougherty, editors, *Microarrays: Optical Technologies and Informatics*, volume 4266 of *Proceedings of SPIE*, pages 141–152, Bellingham, WA, May 2001. SPIE-International Society for Optical Engineering.

Y. H. Yang, M. J. Buckley, S. Dudoit, and T. P. Speed. Comparison of methods for image analysis on cDNA microarray data. *Journal of Computational and Graphical Statistics*, 11(1):108–136, 2002.

D. Yekutieli and Y. Benjamini. Resampling-based false discovery rate controlling multiple test procedures for correlated test statistics. *Journal of Statistical Planning and Inference*, 82:171–196, 1999.

Z. Yu and M. J. van der Laan. Construction of counterfactuals and the G-computation formula. Technical Report 122, Division of Biostatistics, University of California, Berkeley, Berkeley, CA 94720-7360, 2002. Available at `www.bepress.com/ucbbiostat/paper122`.