

References

- A. Alexa, J. Rahnenfuhrer, and T. Lengauer. Improved scoring of functional groups from gene expression data by decorrelating GO graph structure. *Bioinformatics*, 22(13):1600–7, 2006.
- R. Balasubramanian, T. LaFramboise, D. Scholtens, et al. A graph theoretic approach to testing associations between disparate sources of functional genomics data. *Bioinformatics*, 20:3353–3362, 2004.
- Y. Benjamini and Y. Hochberg. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *JRSSB*, 57:289–300, 1995.
- B. M. Bolstad, R. A. Irizarry, L. Gautier, and Z. Wu. Preprocessing high-density oligonucleotide arrays. In R. Gentleman, W. Huber, V. Carey, R. Irizarry, and S. Dudoit, editors, *Bioinformatics and Computational Biology Solutions Using R and Bioconductor*. Springer, New York, 2005.
- L. Breiman. Random forests – random features. *Preprint Department of Statistics U.C. Berkeley*, (567), 1999.
- L. Breiman, J. H. Friedman, R. A. Olshen, et al. *Classification and Regression Trees*. Wadsworth, Belmont, CA, 1984.
- J. Brettschneider, F. Collin, B. Bolstad, and T. Speed. Quality assessment for short oligonucleotide arrays. *Technometrics*, in press, 2007.
- V. Carey, J. Gentry, E. Whalen, and R. Gentleman. Network structures and algorithms in Bioconductor. *Bioinformatics*, 21:135–136, 2005.
- 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*, 103:2771–2778, 2004.
- S. Chiaretti, X. Li, R. Gentleman, A. Vitale, K. S. Wang, F. Mandelli, R. Foa, and J. Ritz. Gene expression profiles of b-lineage adult acute lymphocytic leukemia reveal genetic patterns that identify lineage derivation and distinct mechanisms of transformation. *Clinical Cancer Research*, 11:7209–7219, 2005.
- R. Cho, M. Campbell, E. Winzeler, et al. A genome-wide transcriptional analysis of the mitotic cell cycle. *Molecular Cell*, 2:65–73, 1998.
- S. Dudoit, Y. H. Yang, T. P. Speed, and M. J. Callow. Statistical methods for identifying differentially expressed genes in replicated cDNA microarray experiments. *Statistica Sinica*, 12:111–139, 2002.

- B. P. Durbin, J. S. Hardin, D. M. Hawkins, and D. M. Rocke. A variance-stabilizing transformation for gene-expression microarray data. *Bioinformatics*, 18 Suppl 1:105–110, 2002.
- ENCODE Project Consortium, E. Birney, J. A. Stamatoyannopoulos, A. Dutta, et al. Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. *Nature*, 447:799–816, 2007.
- S. Falcon and R. Gentleman. Using GStats to test gene lists for GO term association. *Bioinformatics*, 23(2):257–8, 2007.
- R. D. Finn, J. Mistry, B. Schuster-Bockler, S. Griffiths-Jones, V. Hollich, T. Lassmann, S. Moxon, M. Marshall, A. Khanna, R. Durbin, S. R. Eddy, E. L. L. Sonnhammer, and A. Bateman. Pfam: Clans, Web tools and services. *Nucleic Acids Res*, 34(Database issue):247–251, Jan 2006.
- P. Flicek, B. L. Aken, K. Beal, B. Ballester, M. Caccamo, Y. Chen, L. Clarke, G. Coates, F. Cunningham, T. Cutts, T. Down, S. C. Dyer, T. Eyre, S. Fitzgerald, J. Fernandez-Banet, S. Gräf, S. Haider, M. Hammond, R. Holland, K. L. Howe, K. Howe, N. Johnson, A. Jenkinson, A. Kähäri, D. Keefe, F. Kokocinski, E. Kulesha, D. Lawson, I. Longden, K. Megy, P. Meidl, B. Overduin, A. Parker, B. Pritchard, A. Prlic, S. Rice, D. Rios, M. Schuster, I. Sealy, G. Slater, D. Smedley, G. Spudich, S. Trevanion, A. J. Vilella, J. Vogel, S. White, M. Wood, E. Birney, T. Cox, V. Curwen, R. Durbin, X. M. Fernandez-Suarez, J. Herrero, T. J. Hubbard, A. Kasprzyk, G. Proctor, J. Smith, A. Ureta-Vidal, and S. Searle. Ensembl 2008. *Nucleic Acids Res*, 2007.
- E. R. Gansner and S. C. North. An open graph visualization system and its applications to software engineering. *Software Practice and Experience*, 30:1203–1233, 1999.
- H. Ge, Z. Liu, G. M. Church, and M. Vidal. Correlation between transcriptome and interactome mapping data from *Saccharomyces cerevisiae*. *Nature Genetics*, 29:482–486, 2001.
- R. Gentleman, W. Huber, V. Carey, R. Irizarry, and S. Dudoit, editors. *Bioinformatics and Computational Biology Solutions Using R and Bioconductor*. Springer, New York, 2005a.
- R. Gentleman, M. Ruschhaupt, and W. Huber. On the synthesis of microarray experiments. *Journal de la Société Françaises de Statistique*, 146 (1-2):173–194, 2005b.
- R. C. Gentleman, V. J. Carey, D. M. Bates, et al. Bioconductor: Open software development for computational biology and bioinformatics. *Genome Biology*, 5: R80, 2004.
- A. D. Gordon. *Classification*. Chapman & Hall CRC, Boca Raton, FL, 2nd edition, 1999.
- T. Hastie, R. Tibshirani, and J. H. Friedman. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer-Verlag, New York, 2001.

- W. Huber, A. von Heydebreck, H. Sültmann, A. Poustka, and M. Vingron. Variance stabilization applied to microarray data calibration and to quantification of differential expression. *Bioinformatics*, 18:96–104, 2002.
- W. Huber, A. von Heydebreck, H. Sültmann, A. Poustka, and M. Vingron. Parameter estimation for the calibration and variance stabilization of microarray data. *Statistical Applications in Genetics and Molecular Biology*, 2(1), 2003.
- W. Huber, A. von Heydebreck, and M. Vingron. *Encyclopedia of Genetics, Genomics, Proteomics and Bioinformatics*, chapter Error models for microarray intensities. Wiley, New York, 2005.
- R. A. Irizarry, B. Hobbs, F. Collin, Y. D. Beazer-Barclay, K. J. Antonellis, U. Scherf, and T. P. Speed. Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics*, 4:249–264, 2003.
- Z. Jiang and R. Gentleman. Extensions to gene set enrichment. *Bioinformatics*, 23:306–313, 2007.
- M. Kanehisa and S. Goto. KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res*, 28:27–30, 2000.
- M. Kanehisa, S. Goto, M. Hattori, K. F. Aoki-Kinoshita, M. Itoh, S. Kawashima, T. Katayama, M. Araki, and M. Hirakawa. From genomics to chemical genomics: new developments in KEGG. *Nucleic Acids Res*, 34(Database issue):354–357, Jan 2006.
- L. Kaufman and P. J. Rousseeuw. *Finding Groups in Data*. Wiley, New York, 1990.
- T. Kohonen. *Self Organizing Maps*. Springer, Berlin, 1995.
- S. K. Kummerfeld and S. A. Teichmann. DBD: A transcription factor prediction database. *Nucleic Acids Res*, 34:D74–D81, 2006.
- A. Liaw and M. Wiener. Classification and regression by randomforest. *R News*, 2(3):18–22, 2002. URL <http://CRAN.R-project.org/doc/Rnews>.
- I. Lönnstedt and T. Speed. Replicated microarray data. *Statistica Sinica*, 12: 31–46, 2002.
- D. Maglott, J. Ostell, K. D. Pruitt, and T. Tatusova. Entrez Gene: gene-centered information at NCBI. *Nucleic Acids Res*, 35(Database issue):26–31, 2007.
- M. McGee and Z. Chen. Parameter estimation for the exponential-normal convolution model for background correction of Affymetrix GeneChip data. *Statistical Applications in Genetics and Molecular Biology*, 5(1):Article 24, 2006.
- N. J. Mulder, R. Apweiler, T. K. Attwood, A. Bairoch, A. Bateman, D. Binns, P. Bork, V. Buillard, L. Cerutti, R. Copley, E. Courcelle, U. Das, L. Daugherty, M. Dibley, R. Finn, W. Fleischmann, J. Gough, D. Haft, N. Hulo, S. Hunter, D. Kahn, A. Kanapin, A. Kejariwal, A. Labarga, P. S. Langendijk-Genevaux, D. Lonsdale, R. Lopez, I. Letunic, M. Madera, J. Maslen, C. McAnulla, J. McDowall, J. Mistry, A. Mitchell, A. N. Nikolskaya, S. Orchard, C. Orengo, R. Petryszak, J. D. Selengut, C. J. Sigris, P. D. Thomas, F. Valentin, D. Wilson,

- C. H. Wu, and C. Yeats. New developments in the interpro database. *Nucleic Acids Res*, 35(Database issue):224–228, 2007.
- P. Murrell. *R Graphics*. Chapman and Hall, Boca Raton, FL, 2005.
- M. S. Pepe, G. Longton, G. L. Anderson, and M. Schummer. Selecting differentially expressed genes from microarray experiments. *Biometrics*, 59(1):133–142, Mar 2003.
- K. S. Pollard and M. J. van der Laan. Cluster analysis of genomic data. In R. Gentleman, W. Huber, V. Carey, R. Irizarry, and S. Dudoit, editors, *Bioinformatics and Computational Biology Solutions Using R and Bioconductor*. Springer, New York, 2005.
- K. S. Pollard, S. Dudoit, and M. J. van der Laan. Multiple testing procedures: The multtest package and applications to genomics. In R. Gentleman, W. Huber, V. Carey, R. Irizarry, and S. Dudoit, editors, *Bioinformatics and Computational Biology Solutions Using R and Bioconductor*. Springer, New York, 2005.
- R-Foundation. Introduction to R, 2007. URL <http://cran.r-project.org/manuals/R-intro.html>.
- B. D. Ripley. *Pattern Recognition and Neural Networks*. Cambridge University Press, Cambridge, 1996.
- M. E. Ritchie, J. Silver, A. Oshlack, M. Holmes, D. Diyagama, A. Holloway, and G. K. Smyth. A comparison of background correction methods for two-colour microarrays. *Bioinformatics*, 23:2700–2707, 2007.
- D. M. Rocke and B. Durbin. Approximate variance-stabilizing transformations for gene-expression microarray data. *Bioinformatics*, 19:966–972, 2003.
- P. J. Rousseuw and A. M. Leroy. *Robust Regression and Outlier Detection*. Wiley, New York, 1987.
- A. Schroeder, O. Mueller, S. Stocker, R. Salowsky, M. Leiber, M. Gassmann, S. Lightfoot, W. Menzel, M. Granzow, and T. Ragg. The RIN: an RNA integrity number for assigning integrity values to RNA measurements. *BMC Mol Biol*, 7: 3, 2006.
- J. G. Siek, L.-Q. Lee, and A. Lumsdaine. *The Boost Graph Library*. Addison Wesley, Boston, 2002.
- G. Smyth. Linear models and empirical Bayes methods for assessing differential expression in microarray experiments. *Statistical Applications in Genetics and Molecular Biology*, 3:Article 3, 2004.
- P. H. A. Sneath and R. R. Sokal. *Numerical Taxonomy: The Principles and Practice of Numerical Classification*. Freeman, San Francisco, 1973.
- A. Subramanian, P. Tamayo, V. K. Mootha, et al. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. of the U.S.A.*, 102(43):15545–15550, 2005.
- H. Sültmann, A. von Heydebreck, W. Huber, R. Künler, A. Bunes, M. Vogt, B. Gunawan, M. Vingron, L. Füzesi, and A. Poustka. Gene expression in kidney

- cancer is associated with cytogenetic abnormalities, metastasis formation, and patient survival. *Clinical Cancer Research*, 11:646–655, 2005.
- The Gene Ontology Consortium. Gene Ontology: tool for the unification of biology. *Nature Genetics*, 25:25–29, 2000.
- L. Tian, S. A. Greenberg, S. W. Kong, et al. Discovering statistically significant pathways in expression profiling studies. *Proc. Natl. Acad. Sci. of the U.S.A.*, 102(38):13544–13549, 2005.
- V. Tusher, R. Tibshirani, and G. Chu. Significance analysis of microarrays applied to the ionizing radiation response. *Proc. Natl. Acad. Sci. of the U.S.A.*, 98: 5116–5121, 2001.
- UniProt. The universal protein resource (UniProt). *Nucleic Acids Res*, 35 (Database issue):193–197, 2007.
- W. N. Venables and B. D. Ripley. *Modern Applied Statistics with S (4e)*. Springer, New York, 2002.
- A. von Heydebreck, W. Huber, and R. Gentleman. Differential expression with the Bioconductor project. In *Encyclopedia of Genetics, Genomics, Proteomics and Bioinformatics*. Wiley, New York, 2004.
- P. Westfall and S. Young. *Resampling-Based Multiple Testing: Examples and Methods for p-Value Adjustment*. Wiley, New York, 1993.
- D. L. Wheeler, T. Barrett, D. A. Benson, S. H. Bryant, K. Canese, V. Chetvernin, D. M. Church, M. DiCuccio, R. Edgar, S. Federhen, L. Y. Geer, Y. Kapustin, O. Khovayko, D. Landsman, D. J. Lipman, T. L. Madden, D. R. Maglott, J. Ostell, V. Miller, K. D. Pruitt, G. D. Schuler, E. Sequeira, S. T. Sherry, K. Sirotkin, A. Souvorov, G. Starchenko, R. L. Tatusov, T. A. Tatusova, L. Wagner, and E. Yaschenko. Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res*, 35(Database issue):5–12, 2007.

Index

- ?(), 5
- [], 223
- [[(), 116, 223

- A-value, 52, 70, 80
- `aaf.handler()`, 39
- `abline()`, 21
- acute lymphoblastic leukemia (ALL),
1, 90, 91, 96, 207
- Affymetrix, 25, 26, 28, 34, 40, 41
- `affyQAResult()`, 28
- Agilent, 48
- `agopen()`, 175
- ALL1/AF4, 2, 4, 103, 104, 108, 135
- annotation, 38, 39, 103, 104, 106–109,
112–117, 119
- `apply()`, 9, 10
- `apropos()`, 5
- attribute, 114, 115

- background correction, 32, 40, 42, 63,
66
- biased, 68
- background-correction, 41–44, 65, 66,
68, 73, 75, 79
- biased, 68
- barplot, 259
- `basecontent()`, 21
- BCR/ABL, 2, 4, 83, 85, 86, 91, 99,
100, 103, 104, 108, 123, 127,
128, 135, 138, 141, 196, 197,
199, 207, 210, 212, 245, 249,
261
- Benjamini and Hochberg, 87
- between-array normalization, 32
- `biocLite()`, 7
- BioMart, 113, 115
- biplot, 156, 256

- Bonferroni correction, 38
- Boost graph library, 173
- boxplot, 22, 53, 54, 195
- `boxplot()`, 23
- `boxplot.stats()`, 23
- `brewer.pal()`, 22
- `browseURL()`, 216

- carbon tetrachloride (CCl4), 48, 54,
58
- CDF, 40, 43
- CEL, 12, 25, 26, 40
- Chi Square Test, 252
- `chisq.test()`, 108
- circo, 175
- class
 - AffyBatch*, 25, 26, 33, 42, 72
 - AnnotatedDataFrame*, 15–19, 28,
49, 54
 - Bimap*, 119
 - character*, 195, 222
 - ChrMapHyperGParams*, 219, 220
 - clusterGraph*, 163
 - data.frame*, 8, 15, 16, 27, 49, 76,
108, 119, 215, 223, 235, 256
 - diana*, 144, 250
 - dist*, 125
 - environment*, 5, 9–11
 - eSet*, 72
 - ExpressionSet*, 1, 5, 11, 12, 15–19,
33, 65, 72, 78, 85, 100, 104,
194, 197, 202, 211, 213
 - factor*, 2, 57, 84, 91, 197
 - GeneSet*, 194, 195
 - GeneSetCollection*, 194
 - GOHyperGParams*, 213, 214, 218
 - GOHyperGResult*, 112, 215, 216
 - GOTerms*, 239

- graph*, 165, 170, 174, 177, 188–190, 258
- graphNEL*, 256
- hclust*, 144, 250
- HyperGResult*, 215
- intactComplex*, 169
- intactGraph*, 171
- intactHyperGraph*, 171
- intactInteraction*, 167
- integer*, 222
- interactionEntry*, 166, 168
- interactions*, 167
- interactors*, 168
- KEGGHyperGParams*, 213, 215, 219, 220
- list*, 8
- logical*, 222
- MArrayLM*, 58
- Mart*, 113, 114
- matrix*, 8, 12, 72, 223
- MIAME*, 17
- NChannelSet*, 54, 55, 70, 72
- organismName*, 167
- PFAMHyperGParams*, 213, 215, 219
- randomForest*, 133
- releaseDate*, 167
- RGList*, 49, 54, 72
- vector*, 8
- vsn*, 76
- class discovery, 137
- classification, 121, 122, 128, 129
- classifier, 121
- CLL, 26, 27, 40
- `closest.top()`, 126, 244
- cluster, 186
- cluster graph, 163, 164
- clustering, 121, 137, 138, 143–149, 151–153, 157, 252–254
- `cmdscale()`, 141, 249
- `coef()`, 74
- `colorRampPalette()`, 22
- `confuMat()`, 127, 131, 243, 245
- confusion matrix, 121, 127, 132, 243, 248
- connected components, 160, 161, 163, 171, 174, 175, 257
- `connectedComp()`, 160, 174
- `cophenetic()`, 145
- cophenetic correlation, 145, 146, 251
- CRAN, 7
- cross-validation, 121, 122, 124, 127, 129–132, 137
 - leave-one-out (LOO), 130
 - ten-fold, 131
- `cutree()`, 145, 250
-
- `daisy()`, 125, 139, 241
- dendrogram, 138, 142–146
- density estimate, 53
- density plot, 22
- diagonal linear discriminant analysis (DLDA), 127, 129
- diameter, 162, 258
- `diana()`, 144, 255
- `dim()`, 33
- dimension reduction, 137
- Dimethylsulfoxid (DMSO), 58
- dimethylsulfoxid (DMSO), 48
- DIP, 166
- directed acyclic graph (DAG), 110, 216
- `dist()`, 125, 139, 241
- `dist2()`, 29
- distance, 121, 122, 124–126, 137, 139–142, 144–147, 150, 241, 242, 244, 250, 251
 - correlation, 125, 241
 - cosangle, 152
 - Euclidean, 124, 128, 241, 242
 - Kullback–Leibler, 241
 - Manhattan, 140, 241
 - mutual information, 241
- `dot`, 175, 179
- dot plot, 20
- dye-swap, 48
-
- `eapply()`, 9, 10
- `eBayes()`, 35, 58, 96
- `edgeNames()`, 177
- `edgeRenderInfo()`, 176–178, 257
- empirical Bayes moderation, 35, 96
- empirical cumulative distribution function (ECDF), 23, 79
- `expresso()`, 12
- `exprs()`, 19, 33

- F-test, 93
- `factor()`, 3
- false discovery rate (FDR), 38, 87, 95
- familywise error rate (FWER), 95
- feature, 121–125, 128–133, 137, 242, 245
- feature selection, 121, 122, 124, 130, 131, 133
- feature space, 122
- `featureNames()`, 18
- filter, 114, 115
- `find()`, 5
- `findEGs()`, 118
- Fisher's exact test, 108, 109, 208
- `fisher.test()`, 108
- fold-change, 63, 68, 79, 80, 90, 93
- `fs.absT()`, 131
- `fsHistory()`, 131
- function, 9

- GC-content, 21, 22
- Gene Ontology (GO), 106, 107, 109–112, 114, 116, 117, 119, 139, 193, 194, 209–216, 218–220
- gene set enrichment analysis (GSEA), 109, 111, 193, 194, 200, 209
- gene universe, 209, 210, 212–214
- Genepix, 49
- generalized log-ratio, 68, 72, 79
- generalized logarithm (glog), 42, 55, 57, 63, 65, 68, 72, 73, 79, 80
- `GeneSetCollection()`, 197
- `get()`, 116
- `getBM()`, 115
- `getGene()`, 114
- `getGo()`, 114
- `getHomolog()`, 114
- `getMA()`, 76
- `getSequence()`, 114
- `getSequences()`, 114
- `getSNP()`, 114
- `getwd()`, 25
- `GO2heatmap()`, 195
- GOA, 193
- `GOmplot()`, 195
- graph, 159–165, 170, 171, 173–177, 179–188, 190, 257, 258
 - directed, 175, 177, 180
 - undirected, 175, 177
- graph layout, 173–177, 179, 181–187, 189
- graph rendering, 173–177, 179, 181, 182, 189
- `graph.par()`, 176, 177, 179
- Graphviz, 173–175, 177, 179, 183, 186
- `grep()`, 10
- `gseattperm()`, 200
- `gsub()`, 11

- `hclust()`, 144
- `head()`, 113
- heatmap, 29, 92, 125, 126, 140, 149, 195, 202, 242, 243
- help, 5, 6
- `help.search()`, 6
- HG-U133A GeneChip, 117
- HG-U95Av2 GeneChip, 9, 17, 21, 26, 32, 34
- `hgu133a_dbconn()`, 115
- `hgu95av2_dbschema()`, 118
- hierarchical clustering, 137, 138, 142, 144–146, 151, 154, 251
 - agglomerative, 138, 143, 144
 - divisive, 138, 143, 144
- `hopach()`, 151
- hopach clustering, 152
- `htmlReport()`, 216
- HUGO, 168
- Hypergeometric distribution, 208
- Hypergeometric test, 193, 194, 207, 210, 212–214, 217, 219
 - conditional, 218, 219
- `hyperGTest()`, 110, 207, 213–216, 218, 220

- image plot, 50, 51
- `imageMap()`, 188, 190
- `importance()`, 133
- `incidence()`, 196
- incidence matrix, 196–198, 203
- index vector, 8, 222
- induced GO graph, 110, 112
- `install.packages()`, 7
- IntAct, 159, 165–169, 171
- `intactXML2Graph()`, 165, 170, 171
- InterPro, 115

- `intersection()`, 164, 258
- IQR, 3, 84, 103, 104, 124, 212
- `IQR()`, 211

- `justRMA()`, 12
- `justvsn()`, 42, 54, 57, 70, 72, 78, 82

- k nearest neighbors (KNN), 127–130, 132, 133, 136, 246, 247, 249
- k-means clustering, 137, 146–148, 151, 254, 255
- KEGG, 180, 181, 183, 186, 193, 194, 196, 197, 202, 213, 219, 220
- `KEGG2heatmap()`, 195, 199
- `KEGGmnpplot()`, 195
- `kmeans()`, 146, 147
- `knn()`, 127, 131, 244
- `knn1()`, 149, 151
- `knnI()`, 127, 131

- `lapply()`, 9, 213
- `layoutGraph()`, 161, 174, 175, 177, 179, 183, 186, 257
- least trimmed sum of squares (LTS), 73, 76, 77
- `levelplot()`, 29
- library, 7
- `library()`, 7
- linear discriminant analysis (LDA), 127, 129
- linear model, 38, 58, 95, 203
- `list()`, 223
- `listAttributes()`, 115
- `listDatasets()`, 113
- `listFilters()`, 114
- `listMarts()`, 113
- `lmFit()`, 58, 95
- log fold-change, 34, 35
- log-ratio, 34, 50, 60, 63–66, 68, 80

- M-value, 52, 58, 70, 72
- MA-plot, 50, 52, 55, 57, 71, 79, 80
- machine learning, 121, 122, 124–131, 139, 140, 243
 - supervised, 121, 122, 124, 127, 129, 135, 137
 - unsupervised, 121, 137, 142, 157
- Mahalanobis transformation, 124
- mailing list, 6
- Mann–Whitney test, 90
- `MAPAmat()`, 202, 263
- `mapply()`, 213
- `matplot()`, 230
- maxT procedure, 95
- `meanSdPlot()`, 74, 91
- median split silhouette (MSS), 144
- medoid, 147
- `mget()`, 116
- `MIdist()`, 126
- mismatch probe (MM), 40, 41, 44, 45, 228, 230
- `MLearn()`, 127, 130, 132, 243
- `mm()`, 40
- model matrix, 57
- `msscheck()`, 144
- `mssEst()`, 144
- `mt.maxT()`, 94, 95
- `mt.raw2adjp()`, 87
- `multidensity()`, 22, 54, 66
- multidimensional scaling, 141, 142, 175
- `multiecdf()`, 23
- multiple testing, 34, 38, 60, 83, 87, 89, 94, 103, 111, 233
- `myExtract()`, 11

- `names()`, 251
- `nchar()`, 182
- neato, 172, 175, 179
- `new()`, 17, 163
- `nodePerm()`, 164, 258
- `nodeRenderInfo()`, 176–178, 182, 257
- `nodes()`, 160, 177
- nonspecific lering, 86
- nonspecific filtering, 1, 3, 34, 42, 83–87, 89, 90, 103, 123, 124, 128, 135, 197, 209–214, 229, 233
- normalization, 40, 47, 54, 89
- `normalizeBetweenArrays()`, 57
- `normalizeWithinArrays()`, 52, 57
- `nrssel.pAUC()`, 102
- `nrssel.ttest()`, 102
- `nsFilter()`, 3, 4, 34, 85, 103, 108, 123, 197, 211, 235, 265
- NUSE, 30, 31

- openHtmlPage(), 190
- openVignette(), 6
- ordinary least squares, 76

- p.adjust(), 38
- package, 7, 8, 221
 - affy, 12, 25, 32
 - affyPLM, 12, 30, 33
 - affyQCReport, 28
 - ALL, 2, 138, 193, 210
 - annaffy, 39, 106
 - annotate, 38, 110
 - AnnotationDbi, 115, 119
 - arrayMagic, 12
 - arrayQualityMetrics, 28
 - base, 49
 - Biobase, 6, 7, 11, 15, 28
 - BiocCaseStudies, 101
 - biocGraph, 188
 - bioDist, 125, 126, 139, 242, 244, 250
 - biomaRt, 113–115
 - Category, 85, 193, 200, 202, 207, 209, 213, 219
 - CCl4, 48, 65
 - class, 127, 130, 148, 149
 - CLL, 21, 26, 27
 - cluster, 125, 139, 144, 147, 152, 241
 - genefilter, 3, 34, 84, 92, 99, 103, 123, 197, 198, 211
 - genepLOT, 22, 66
 - GlobalAncova, 194
 - GO.db, 239
 - GOstats, 110, 111, 152, 194, 207, 209
 - goTools, 110
 - graph, 7, 159, 173, 186
 - graphviz, 181
 - GSEABase, 193–197, 203
 - hgu133a.db, 115
 - hgu95av2, 115
 - hgu95av2.db, 106, 115
 - hopach, 137, 142, 151
 - kohonen, 148
 - lattice, 76
 - limma, 12, 35, 38, 47, 52, 55, 57, 58, 89, 95, 213
 - MASS, 141
 - matchprobes, 21
 - MLInterfaces, 121, 122, 127, 130, 132
 - multtest, 38, 87, 94, 95
 - oligo, 12
 - org.Hs.eg.db, 188
 - PGSEA, 194
 - randomForest, 132
 - RBGL, 159–161
 - RColorBrewer, 22, 125, 140
 - RGL, 174
 - rgl, 141
 - Rgraphviz, 159, 161, 173–176, 178, 179, 184, 185, 260
 - Rintact, 165, 166
 - ROC, 100
 - ScISI, 165
 - sigPathways, 194
 - simpleaffy, 28, 29
 - som, 148
 - topGO, 110, 111, 194
 - vsn, 54, 56, 57, 63, 65, 68, 70, 72, 81, 82
 - xtable, 7, 116
 - YEAST, 214
 - yeastExpData, 160, 174
- pairs plot, 155, 156
- pam(), 147, 252
- partial area under the curve (pAUC), 99, 100, 102
- partitioning, 138, 146, 151
- partitioning around mediods (PAM), 147, 148, 152, 252
- paste(), 9
- pData(), 16
- Pearson's χ^2 test, 108
- perfect match probe (PM), 40–45, 228–230
- permutation distribution, 259
- permutation test, 90, 200, 201
- PFAM, 112, 114, 115, 213, 219, 220
- phenoData(), 19
- Philadelphia chromosome, 2, 91
- plot(), 20
- plotMA(), 52, 231
- pm(), 40
- pm.abstGrep(), 112
- pm.getabst(), 112
- pnorm(), 262

- `points()`, 36
- `prcomp()`, 154
- `predict()`, 72
- preprocessing, 32, 33, 40, 63
- principal components, 154
- Prosite, 112, 115
- PSI-MI 2.5 file, 165, 166, 168, 170, 171
- `psi25complex()`, 165, 169, 170
- `psi25interaction()`, 165, 166, 169, 170
- PubMed, 112, 188, 190
- q-q plot, 198
- q-q plot, 195
- QA (quality assessment), 28
- `qc()`, 28
- QC (quality control), 28, 29
- R-FAQ, 6
- random forests, 132–136, 248
- `rbind()`, 76
- `read.maimages()`, 47, 49, 50, 55
- `read.table()`, 12, 13, 15
- `ReadAffy()`, 12, 25, 26
- Receiver Operator Characteristic (ROC), 89, 90, 99–101, 124, 234
- reference distribution, 160
- reference normalization, 81, 82
- `renderGraph()`, 161, 175–177, 179, 182, 190, 257
- `renderParameters()`, 260
- reporter summarization, 32, 40, 42, 43
- `require()`, 7
- `resample()`, 101
- `return()`, 9
- `revmap()`, 119
- RLE, 30, 31
- RMA, 32, 42–44, 229
- `rma()`, 1, 32, 33, 42
- RNA Integrity Number (RIN), 48
- RNA integrity number (RIN), 48
- `rowFtests()`, 93, 198
- `rowMeans()`, 22, 34
- `rowAUCs()`, 101
- `rowQ()`, 84, 93
- `rowSds()`, 84
- `rowttests()`, 34, 85, 92, 93, 99, 101, 104, 198
- `sammon()`, 141, 249
- Sammon mapping, 141, 250
- `sampleNames()`, 16, 19
- `sapply()`, 9, 15
- search path, 5
- Self-organizing maps (SOMs), 148
- self-organizing maps (SOMs), 148
- sensitivity, 99
- `sessionInfo()`, 8, 221
- `setwd()`, 25
- shortest path, 161–163, 258
- shorth, 84
- `shorth()`, 84
- shrinkage estimator, 63, 68, 80
- `silcheck()`, 144
- `silhouette()`, 152–154
- silhouette plot, 142, 152, 153
- single-linkage clustering, 251
- `smoothScatter()`, 52, 231
- SNP, 114
- SOM(), 148, 149
- `sp.between()`, 257
- `spearman.dist()`, 242
- specificity, 99
- `split()`, 163
- spot shape descriptors, 47
- SQLite, 115
- standardization, 124
- `standardize()`, 124
- Student's t-distribution, 34, 36
- `subGraph()`, 161, 174, 186
- subgraph, 174, 185–187
- `summary()`, 215, 219
- support vector machines, 122
- `svd()`, 154
- `t()`, 125
- t-distribution, 194
- t-statistic, 34–36, 43, 44, 86, 104, 105, 124, 128, 131, 142, 193, 194, 200, 203, 205, 250
 - moderated, 35–37, 58, 60, 95
- t-test, 38, 85, 87, 90, 93–95, 97, 98, 101, 102, 124, 128, 133, 210, 212, 213, 234

- moderated, 59, 97, 98
- `table()`, 145, 149, 235, 251
- target file, 49
- `termGraphs()`, 216
- test set, 121, 127, 130, 132, 134, 135, 246, 249
- `threestep()`, 33
- tooltip, 187
- `topTable()`, 38, 96
- `toTable()`, 119
- training set, 121, 127–130, 132, 134, 135, 137, 246, 247
- two-color array, 47, 48
- `twopi`, 172, 175
- `update.packages()`, 7, 8
- variance–bias trade-off, 79, 80
- `varLabels()`, 19
- `varMetadata()`, 16
- vignette, 6
- visualization, 5, 47, 141, 145
- volcano plot, 35–37, 60, 61, 93
- VSN, 42–44, 54, 70, 72–75, 77, 78, 80, 81, 229
- `vsn()`, 72
- `vsn2()`, 72, 74, 81
- `vsnrma()`, 42
- Welch statistic, 94
- `xval()`, 131
- `xvalSpec()`, 131