

7 References

- Acton, T. B., H. Zhong, and A. K. Vershon. 1997. DNA-binding specificity of Mcm1: operator mutations that alter DNA-bending and transcriptional activities by a MADS box protein. *Mol Cell Biol* **17**:1881-1889.
- Aerts, S., G. Thijs, B. Coessens, M. Staes, Y. Moreau, and B. De Moor. 2003. Toucan: deciphering the cis-regulatory logic of coregulated genes. *Nucleic Acids Res* **31**:1753-1764.
- Ahituv, N., S. Prabhakar, F. Poulin, E. M. Rubin, and O. Couronne. 2005. Mapping cis-regulatory domains in the human genome using multi-species conservation of synteny. *Hum Mol Genet* **14**:3057-3063.
- Al-Shahrour, F., R. Diaz-Uriarte, and J. Dopazo. 2004. FatiGO: a web tool for finding significant associations of Gene Ontology terms with groups of genes. *Bioinformatics* **20**:578-580.
- Alam, J., and J. L. Cook. 1990. Reporter genes: application to the study of mammalian gene transcription. *Anal Biochem* **188**:245-254.
- Alves, G., A. Tatro, and T. Fanning. 1996. Differential methylation of human LINE-1 retrotransposons in malignant cells. *Gene* **176**:39-44.
- Antequera, F. 2003. Structure, function and evolution of CpG island promoters. *Cell Mol Life Sci* **60**:1647-1658.
- Antequera, F., and A. Bird. 1993. Number of CpG islands and genes in human and mouse. *Proc Natl Acad Sci U S A* **90**:11995-11999.
- Bailey, T. L., and C. Elkan. 1994. Fitting a mixture model by expectation maximization to discover motifs in biopolymers. Pp. 28-36. International Conference on Intelligent Systems for Molecular Biology. AAAI Press, Menlo Park, California.
- Bajic, V. B., and S. H. Seah. 2003a. Dragon Gene Start Finder identifies approximate locations of the 5' ends of genes. *Nucleic Acids Res* **31**:3560-3563.
- Bajic, V. B., and S. H. Seah. 2003b. Dragon gene start finder: an advanced system for finding approximate locations of the start of gene transcriptional units. *Genome Res* **13**:1923-1929.
- Bajic, V. B., S. H. Seah, A. Chong, S. P. Krishnan, J. L. Koh, and V. Brusic. 2003. Computer model for recognition of functional transcription start sites in RNA polymerase II promoters of vertebrates. *J Mol Graph Model* **21**:323-332.
- Bajic, V. B., S. L. Tan, Y. Suzuki, and S. Sugano. 2004. Promoter prediction analysis on the whole human genome. *Nat Biotechnol* **22**:1467-1473.
- Bamshad, M. J., S. Mummidi, E. Gonzalez, S. S. Ahuja, D. M. Dunn, W. S. Watkins, S. Wooding, A. C. Stone, L. B. Jorde, R. B. Weiss, and S. K. Ahuja. 2002. A strong signature of balancing selection in the 5' cis-regulatory region of CCR5. *Proc Natl Acad Sci U S A* **99**:10539-10544.
- Barrett, J. C., B. Fry, J. Maller, and M. J. Daly. 2005. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* **21**:263-265.
- Barrett, T. B., R. L. Hauger, J. L. Kennedy, A. D. Sadovnick, R. A. Remick, P. E. Keck, S. L. McElroy, M. Alexander, S. H. Shaw, and J. R. Kelsoe. 2003. Evidence that a single nucleotide polymorphism in the promoter of the G protein receptor kinase 3 gene is associated with bipolar disorder. *Mol Psychiatry* **8**:546-557.
- Baylin, S. B. 2005. DNA methylation and gene silencing in cancer. *Nat Clin Pract Oncol* **2 Suppl 1**:S4-11.

- Belanger, H., P. Beaulieu, C. Moreau, D. Labuda, T. J. Hudson, and D. Sinnett. 2005. Functional promoter SNPs in cell cycle checkpoint genes. *Hum Mol Genet* **14**:2641-2648.
- Bell, A. C., and G. Felsenfeld. 2000. Methylation of a CTCF-dependent boundary controls imprinted expression of the *Igf2* gene. *Nature* **405**:482-485.
- Bell, A. C., A. G. West, and G. Felsenfeld. 2001. Insulators and boundaries: versatile regulatory elements in the eukaryotic. *Science* **291**:447-450.
- Belting, H. G., C. S. Shashikant, and F. H. Ruddle. 1998. Modification of expression and cis-regulation of *Hoxc8* in the evolution of diverged axial morphology. *Proc Natl Acad Sci U S A* **95**:2355-2360.
- Benjamini, Y., and Y. Hochberg. 1995. Controlling the false discovery rate - a practical and powerful approach to multiple testing. *J Roy Stat Soc B Met* **57**:289-300.
- Bernard, P., and M. Couturier. 1992. Cell killing by the F plasmid CcdB protein involves poisoning of DNA-topoisomerase II complexes. *J Mol Biol* **226**:735-745.
- Blackwood, E. M., and J. T. Kadonaga. 1998. Going the distance: a current view of enhancer action. *Science* **281**:60-63.
- Bray, N. J., P. R. Buckland, M. J. Owen, and M. C. O'Donovan. 2003. Cis-acting variation in the expression of a high proportion of genes in human brain. *Hum Genet* **113**:149-153.
- Brem, R. B., G. Yvert, R. Clinton, and L. Kruglyak. 2002. Genetic dissection of transcriptional regulation in budding yeast. *Science* **296**:752-755.
- Brown, R. P., and M. E. Feder. 2005. Reverse transcriptional profiling: non-correspondence of transcript level variation and proximal promoter polymorphism. *BMC Genomics* **6**:110.
- Bucher, P. 1990. Weight matrix descriptions of four eukaryotic RNA polymerase II promoter elements derived from 502 unrelated promoter sequences. *J Mol Biol* **212**:563-578.
- Buckland, P. R. 2006. The importance and identification of regulatory polymorphisms and their mechanisms of action. *Biochim Biophys Acta* **1762**:17-28.
- Buckland, P. R., S. L. Coleman, B. Hoogendoorn, C. Guy, S. K. Smith, and M. C. O'Donovan. 2004a. A high proportion of chromosome 21 promoter polymorphisms influence transcriptional activity. *Gene Expr* **11**:233-239.
- Buckland, P. R., B. Hoogendoorn, S. L. Coleman, C. A. Guy, S. K. Smith, and M. C. O'Donovan. 2005. Strong bias in the location of functional promoter polymorphisms. *Hum Mutat* **26**:214-223.
- Buckland, P. R., B. Hoogendoorn, C. A. Guy, S. L. Coleman, S. K. Smith, J. D. Buxbaum, V. Haroutunian, and M. C. O'Donovan. 2004b. A high proportion of polymorphisms in the promoters of brain expressed genes influences transcriptional activity. *Biochim Biophys Acta* **1690**:238-249.
- Bulger, M., T. Sawado, D. Schubeler, and M. Groudine. 2002. ChIPs of the beta-globin locus: unraveling gene regulation within an active domain. *Curr Opin Genet Dev* **12**:170-177.
- Bulyk, M. L., P. L. Johnson, and G. M. Church. 2002. Nucleotides of transcription factor binding sites exert interdependent effects on the binding affinities of transcription factors. *Nucleic Acids Res* **30**:1255-1261.
- Burke, T. W., and J. T. Kadonaga. 1996. *Drosophila* TFIID binds to a conserved downstream basal promoter element that is present in many TATA-box-deficient promoters. *Genes Dev* **10**:711-724.

- Butler, J. E., and J. T. Kadonaga. 2001. Enhancer-promoter specificity mediated by DPE or TATA core promoter motifs. *Genes Dev* **15**:2515-2519.
- Caceres, M., J. Lachuer, M. A. Zapala, J. C. Redmond, L. Kudo, D. H. Geschwind, D. J. Lockhart, T. M. Preuss, and C. Barlow. 2003. Elevated gene expression levels distinguish human from non-human primate brains. *Proc Natl Acad Sci U S A* **100**:13030-13035.
- Cajiao, I., A. Zhang, E. J. Yoo, N. E. Cooke, and S. A. Liebhaber. 2004. Bystander gene activation by a locus control region. *Embo J* **23**:3854-3863.
- Carninci, P., A. Sandelin, B. Lenhard, S. Katayama, K. Shimokawa, J. Ponjavic, C. A. Semple, M. S. Taylor, P. G. Engstrom, M. C. Frith, A. R. Forrest, W. B. Alkema, S. L. Tan, C. Plessy, R. Kodzius, T. Ravasi, T. Kasukawa, S. Fukuda, M. Kanamori-Katayama, Y. Kitazume, H. Kawaji, C. Kai, M. Nakamura, H. Konno, K. Nakano, S. Mottagui-Tabar, P. Arner, A. Chesi, S. Gustincich, F. Persichetti, H. Suzuki, S. M. Grimmond, C. A. Wells, V. Orlando, C. Wahlestedt, E. T. Liu, M. Harbers, J. Kawai, V. B. Bajic, D. A. Hume, and Y. Hayashizaki. 2006. Genome-wide analysis of mammalian promoter architecture and evolution. *Nat Genet*.
- Carroll, S. B. 2000. Endless forms: the evolution of gene regulation and morphological diversity. *Cell* **101**:577-580.
- Carson, S., and M. V. Wiles. 1993. Far upstream regions of class II MHC Ea are necessary for position-independent, copy-dependent expression of Ea transgene. *Nucleic Acids Res* **21**:2065-2072.
- Carter, D., L. Chakalova, C. S. Osborne, Y. F. Dai, and P. Fraser. 2002. Long-range chromatin regulatory interactions in vivo. *Nat Genet* **32**:623-626.
- Cawley, S., 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. 2004. Unbiased mapping of transcription factor binding sites along human chromosomes 21 and 22 points to widespread regulation of noncoding RNAs. *Cell* **116**:499-509.
- Chen, L. 1999. Combinatorial gene regulation by eukaryotic transcription factors. *Curr Opin Struct Biol* **9**:48-55.
- Cheng, J., P. Kapranov, J. Drenkow, S. Dike, S. Brubaker, S. Patel, J. Long, D. Stern, H. Tammana, G. Helt, V. Sementchenko, A. Piccolboni, S. Bekiranov, D. K. Bailey, M. Ganesh, S. Ghosh, I. Bell, D. S. Gerhard, and T. R. Gingeras. 2005. Transcriptional maps of 10 human chromosomes at 5-nucleotide resolution. *Science* **308**:1149-1154.
- Cheung, V. G., L. K. Conlin, T. M. Weber, M. Arcaro, K. Y. Jen, M. Morley, and R. S. Spielman. 2003. Natural variation in human gene expression assessed in lymphoblastoid cells. *Nat Genet* **33**:422-425.
- Cheung, V. G., R. S. Spielman, K. G. Ewens, T. M. Weber, M. Morley, and J. T. Burdick. 2005. Mapping determinants of human gene expression by regional and genome-wide association. *Nature* **437**:1365-1369.
- Chuang, J. H., and H. Li. 2004. Functional bias and spatial organization of genes in mutational hot and cold regions in the human genome. *PLoS Biol* **2**:E29.
- Collins, J. E., M. E. Goward, C. G. Cole, L. J. Smink, E. J. Huckle, S. Knowles, J. M. Bye, D. M. Beare, and I. Dunham. 2003. Reevaluating human gene annotation: a second-generation analysis of chromosome 22. *Genome Res* **13**:27-36.

- Collins, J. E., C. L. Wright, C. A. Edwards, M. P. Davis, J. A. Grinham, C. G. Cole, M. E. Goward, B. Aguado, M. Mallya, Y. Mokrab, E. J. Huckle, D. M. Beare, and I. Dunham. 2004. A genome annotation-driven approach to cloning the human ORFeome. *Genome Biol* **5**:R84.
- Concino, M. F., R. F. Lee, J. P. Merryweather, and R. Weinmann. 1984. The adenovirus major late promoter TATA box and initiation site are both necessary for transcription in vitro. *Nucleic Acids Res* **12**:7423-7433.
- Consortium, I. H. G. S. 2004a. Finishing the euchromatic sequence of the human genome. *Nature* **431**:931-945.
- Consortium, T. C. S. a. A. 2005a. Initial sequence of the chimpanzee genome and comparison with the human genome. *Nature* **437**:69-87.
- Consortium, T. E. P. 2004b. The ENCODE (ENCyclopedia Of DNA Elements) Project. *Science* **306**:636-640.
- Consortium, T. I. H. 2005b. A haplotype map of the human genome. *Nature* **437**:1299-1320.
- Cooper, S. J., N. D. Trinklein, E. D. Anton, L. Nguyen, and R. M. Myers. 2006. Comprehensive analysis of transcriptional promoter structure and function in 1% of the human genome. *Genome Res* **16**:1-10.
- Crawford, D. L., J. A. Segal, and J. L. Barnett. 1999. Evolutionary analysis of TATA-less proximal promoter function. *Mol Biol Evol* **16**:194-207.
- Daborn, P. J., J. L. Yen, M. R. Bogwitz, G. Le Goff, E. Feil, S. Jeffers, N. Tijet, T. Perry, D. Heckel, P. Batterham, R. Feyereisen, T. G. Wilson, and R. H. ffrench-Constant. 2002. A single p450 allele associated with insecticide resistance in *Drosophila*. *Science* **297**:2253-2256.
- Davuluri, R. V., I. Grosse, and M. Q. Zhang. 2001. Computational identification of promoters and first exons in the human genome. *Nat Genet* **29**:412-417.
- Dawson, E., Y. Chen, S. Hunt, L. J. Smink, A. Hunt, K. Rice, S. Livingston, S. Bumpstead, R. Bruskiewich, P. Sham, R. Ganske, M. Adams, K. Kawasaki, N. Shimizu, S. Minoshima, B. Roe, D. Bentley, and I. Dunham. 2001. A SNP resource for human chromosome 22: extracting dense clusters of SNPs from the genomic sequence. *Genome Res* **11**:170-178.
- De Gobbi, M., V. Viprakasit, J. R. Hughes, C. Fisher, V. J. Buckle, H. Ayyub, R. J. Gibbons, D. Vernimmen, Y. Yoshinaga, P. de Jong, J. F. Cheng, E. M. Rubin, W. G. Wood, D. Bowden, and D. R. Higgs. 2006. A regulatory SNP causes a human genetic disease by creating a new transcriptional promoter. *Science* **312**:1215-1217.
- de Souza, O. N., and R. L. Ornstein. 1998. Inherent DNA curvature and flexibility correlate with TATA box functionality. *Biopolymers* **46**:403-415.
- Dekker, J., K. Rippe, M. Dekker, and N. Kleckner. 2002. Capturing chromosome conformation. *Science* **295**:1306-1311.
- Deng, G., G. A. Song, E. Pong, M. Sleisenger, and Y. S. Kim. 2004. Promoter methylation inhibits APC gene expression by causing changes in chromatin conformation and interfering with the binding of transcription factor CCAAT-binding factor. *Cancer Res* **64**:2692-2698.
- Dermitzakis, E. T., and A. G. Clark. 2002. Evolution of transcription factor binding sites in Mammalian gene regulatory regions: conservation and turnover. *Mol Biol Evol* **19**:1114-1121.
- Dermitzakis, E. T., A. Reymond, and S. E. Antonarakis. 2005. Conserved non-genic sequences - an unexpected feature of mammalian genomes. *Nat Rev Genet* **6**:151-157.

- Deutsch, S., R. Lyle, E. T. Dermitzakis, H. Attar, L. Subrahmanyam, C. Gehrig, L. Parand, M. Gagnebin, J. Rougemont, C. V. Jongeneel, and S. E. Antonarakis. 2005. Gene expression variation and expression quantitative trait mapping of human chromosome 21 genes. *Hum Mol Genet* **14**:3741-3749.
- Dierks, P., A. van Ooyen, M. D. Cochran, C. Dobkin, J. Reiser, and C. Weissmann. 1983. Three regions upstream from the cap site are required for efficient and accurate transcription of the rabbit beta-globin gene in mouse 3T6 cells. *Cell* **32**:695-706.
- Down, T. A., and T. J. Hubbard. 2005. NestedMICA: sensitive inference of over-represented motifs in nucleic acid sequence. *Nucleic Acids Res* **33**:1445-1453.
- Down, T. A., and T. J. Hubbard. 2002. Computational detection and location of transcription start sites in mammalian genomic DNA. *Genome Res* **12**:458-461.
- Dunham, I., N. Shimizu, B. A. Roe, S. Chisoe, A. R. Hunt, J. E. Collins, R. Bruskiwich, D. M. Beare, M. Clamp, L. J. Smink, R. Ainscough, J. P. Almeida, A. Babbage, C. Bagguley, J. Bailey, K. Barlow, K. N. Bates, O. Beasley, C. P. Bird, S. Blakey, A. M. Bridgeman, D. Buck, J. Burgess, W. D. Burrill, K. P. O'Brien, and et al. 1999. The DNA sequence of human chromosome 22. *Nature* **402**:489-495.
- Dvir, A., J. W. Conaway, and R. C. Conaway. 2001. Mechanism of transcription initiation and promoter escape by RNA polymerase II. *Curr Opin Genet Dev* **11**:209-214.
- Eisen, M. B., P. T. Spellman, P. O. Brown, and D. Botstein. 1998. Cluster analysis and display of genome-wide expression patterns. *Proc Natl Acad Sci U S A* **95**:14863-14868.
- Elander, N., P. Soderkvist, and K. Fransen. 2006. Matrix metalloproteinase (MMP) - 1, -2, -3 and -9 promoter polymorphisms in colorectal cancer. *Anticancer Res* **26**:791-795.
- Enard, W., P. Khaitovich, J. Klose, S. Zollner, F. Heissig, P. Giavalisco, K. Nieselt-Struwe, E. Muchmore, A. Varki, R. Ravid, G. M. Doxiadis, R. E. Bontrop, and S. Paabo. 2002. Intra- and interspecific variation in primate gene expression patterns. *Science* **296**:340-343.
- Enright, A. J., and C. A. Ouzounis. 2001. BioLayout--an automatic graph layout algorithm for similarity visualization. *Bioinformatics* **17**:853-854.
- Euskirchen, G., T. E. Royce, P. Bertone, R. Martone, J. L. Rinn, F. K. Nelson, F. Sayward, N. M. Luscombe, P. Miller, M. Gerstein, S. Weissman, and M. Snyder. 2004. CREB binds to multiple loci on human chromosome 22. *Mol Cell Biol* **24**:3804-3814.
- Evans, R., J. A. Fairley, and S. G. Roberts. 2001. Activator-mediated disruption of sequence-specific DNA contacts by the general transcription factor TFIIB. *Genes Dev* **15**:2945-2949.
- Evans, R. M. 1988. The steroid and thyroid hormone receptor superfamily. *Science* **240**:889-895.
- Ewing, B., and P. Green. 1998. Base-calling of automated sequencer traces using phred. II. Error probabilities. *Genome Res* **8**:186-194.
- Ewing, B., L. Hillier, M. C. Wendl, and P. Green. 1998. Base-calling of automated sequencer traces using phred. I. Accuracy assessment. *Genome Res* **8**:175-185.

- Fay, J. C., H. L. McCullough, P. D. Sniegowski, and M. B. Eisen. 2004. Population genetic variation in gene expression is associated with phenotypic variation in *Saccharomyces cerevisiae*. *Genome Biol* **5**:R26.
- Feinberg, A. P., and B. Vogelstein. 1983. Hypomethylation of ras oncogenes in primary human cancers. *Biochem Biophys Res Commun* **111**:47-54.
- Fickett, J. W., and A. G. Hatzigeorgiou. 1997. Eukaryotic promoter recognition. *Genome Res* **7**:861-878.
- Filippova, G. N., C. P. Thienes, B. H. Penn, D. H. Cho, Y. J. Hu, J. M. Moore, T. R. Klesert, V. V. Lobanenkova, and S. J. Tapscott. 2001. CTCF-binding sites flank CTG/CAG repeats and form a methylation-sensitive insulator at the DM1 locus. *Nat Genet* **28**:335-343.
- Garcia-Giralt, N., A. Enjuanes, M. Bustamante, L. Mellibovsky, X. Nogues, R. Carreras, A. Diez-Perez, D. Grinberg, and S. Balcells. 2005. In vitro functional assay of alleles and haplotypes of two COL1A1-promoter SNPs. *Bone* **36**:902-908.
- Garcia-Giralt, N., X. Nogues, A. Enjuanes, J. Puig, L. Mellibovsky, A. Bay-Jensen, R. Carreras, S. Balcells, A. Diez-Perez, and D. Grinberg. 2002. Two new single-nucleotide polymorphisms in the COL1A1 upstream regulatory region and their relationship to bone mineral density. *J Bone Miner Res* **17**:384-393.
- Gardiner-Garden, M., and M. Frommer. 1987. CpG islands in vertebrate genomes. *J Mol Biol* **196**:261-282.
- Ge, B., S. Gurd, T. Gaudin, C. Dore, P. Lepage, E. Harmsen, T. J. Hudson, and T. Pastinen. 2005. Survey of allelic expression using EST mining. *Genome Res* **15**:1584-1591.
- Gentleman, R. C., V. J. Carey, D. M. Bates, B. Bolstad, M. Dettling, S. Dudoit, B. Ellis, L. Gautier, Y. Ge, J. Gentry, K. Hornik, T. Hothorn, W. Huber, S. Iacus, R. Irizarry, F. Leisch, C. Li, M. Maechler, A. J. Rossini, G. Sawitzki, C. Smith, G. Smyth, L. Tierney, J. Y. Yang, and J. Zhang. 2004. Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol* **5**:R80.
- Giardine, B., C. Riemer, R. C. Hardison, R. Burhans, L. Elnitski, P. Shah, Y. Zhang, D. Blankenberg, I. Albert, J. Taylor, W. Miller, W. J. Kent, and A. Nekrutenko. 2005. Galaxy: a platform for interactive large-scale genome analysis. *Genome Res* **15**:1451-1455.
- Gibson, A. W., J. C. Edberg, J. Wu, R. G. Westendorp, T. W. Huizinga, and R. P. Kimberly. 2001. Novel single nucleotide polymorphisms in the distal IL-10 promoter affect IL-10 production and enhance the risk of systemic lupus erythematosus. *J Immunol* **166**:3915-3922.
- Gilad, Y., A. Oshlack, G. K. Smyth, T. P. Speed, and K. P. White. 2006. Expression profiling in primates reveals a rapid evolution of human transcription factors. *Nature* **440**:242-245.
- Grosschedl, R., and M. L. Birnstiel. 1980. Identification of regulatory sequences in the prelude sequences of an H2A histone gene by the study of specific deletion mutants in vivo. *Proc Natl Acad Sci U S A* **77**:1432-1436.
- Hain, J., W. D. Reiter, U. Hudepohl, and W. Zillig. 1992. Elements of an archaeal promoter defined by mutational analysis. *Nucleic Acids Res* **20**:5423-5428.
- Hamblin, M. T., and A. Di Rienzo. 2000. Detection of the signature of natural selection in humans: evidence from the Duffy blood group locus. *Am J Hum Genet* **66**:1669-1679.

- Hark, A. T., C. J. Schoenherr, D. J. Katz, R. S. Ingram, J. M. LeVorse, and S. M. Tilghman. 2000. CTCF mediates methylation-sensitive enhancer-blocking activity at the H19/Igf2 locus. *Nature* **405**:486-489.
- Harr, B., and C. Schlotterer. 2006. Comparison of algorithms for the analysis of Affymetrix microarray data as evaluated by co-expression of genes in known operons. *Nucleic Acids Res* **34**:e8.
- Hartl, D. L., and A. G. Clark. 1997. Principles of population genetics. Sinauer, Sunderland, Mass.
- Hashimoto, S., Y. Suzuki, Y. Kasai, K. Morohoshi, T. Yamada, J. Sese, S. Morishita, S. Sugano, and K. Matsushima. 2004. 5'-end SAGE for the analysis of transcriptional start sites. *Nat Biotechnol* **22**:1146-1149.
- Heissig, F., J. Krause, J. Bryk, P. Khaitovich, W. Enard, and S. Paabo. 2005. Functional analysis of human and chimpanzee promoters. *Genome Biol* **6**:R57.
- Herman, J. G., F. Latif, Y. Weng, M. I. Lerman, B. Zbar, S. Liu, D. Samid, D. S. Duan, J. R. Gnarr, W. M. Linehan, and et al. 1994. Silencing of the VHL tumor-suppressor gene by DNA methylation in renal carcinoma. *Proc Natl Acad Sci U S A* **91**:9700-9704.
- Hertz, G. Z., and G. D. Stormo. 1999. Identifying DNA and protein patterns with statistically significant alignments of multiple sequences. *Bioinformatics* **15**:563-577.
- Hinds, D. A., L. L. Stuve, G. B. Nilsen, E. Halperin, E. Eskin, D. G. Ballinger, K. A. Frazer, and D. R. Cox. 2005. Whole-genome patterns of common DNA variation in three human populations. *Science* **307**:1072-1079.
- Ho, Y., F. Elefant, N. Cooke, and S. Liebhaber. 2002. A defined locus control region determinant links chromatin domain acetylation with long-range gene activation. *Mol Cell* **9**:291-302.
- Ho, Y., S. A. Liebhaber, and N. E. Cooke. 2004. Activation of the human GH gene cluster: roles for targeted chromatin modification. *Trends Endocrinol Metab* **15**:40-45.
- Hoffmann, R., T. Seidl, and M. Dugas. 2002. Profound effect of normalization on detection of differentially expressed genes in oligonucleotide microarray data analysis. *Genome Biol* **3**:RESEARCH0033.
- Hoogendoorn, B., S. L. Coleman, C. A. Guy, K. Smith, T. Bowen, P. R. Buckland, and M. C. O'Donovan. 2003. Functional analysis of human promoter polymorphisms. *Hum Mol Genet* **12**:2249-2254.
- Hu, S. L., and J. L. Manley. 1981. DNA sequence required for initiation of transcription in vitro from the major late promoter of adenovirus 2. *Proc Natl Acad Sci U S A* **78**:820-824.
- Huppert, J. L., and S. Balasubramanian. 2005. Prevalence of quadruplexes in the human genome. *Nucleic Acids Res* **33**:2908-2916.
- Huttley, G. A., M. W. Smith, M. Carrington, and S. J. O'Brien. 1999. A scan for linkage disequilibrium across the human genome. *Genetics* **152**:1711-1722.
- Irizarry, R. A., Z. Wu, and H. A. Jaffee. 2006. Comparison of Affymetrix GeneChip expression measures. *Bioinformatics* **22**:789-794.
- Iwama, H., and T. Gojobori. 2004. Highly conserved upstream sequences for transcription factor genes and implications for the regulatory network. *Proc Natl Acad Sci U S A* **101**:17156-17161.

- Javahery, R., A. Khachi, K. Lo, B. Zenzie-Gregory, and S. T. Smale. 1994. DNA sequence requirements for transcriptional initiator activity in mammalian cells. *Mol Cell Biol* **14**:116-127.
- Jin, V. X., G. A. Singer, F. J. Agosto-Perez, S. Liyanarachchi, and R. V. Davuluri. 2006. Genome-wide analysis of core promoter elements from conserved human and mouse orthologous pairs. *BMC Bioinformatics* **7**:114.
- Kanduri, C., V. Pant, D. Loukinov, E. Pugacheva, C. F. Qi, A. Wolffe, R. Ohlsson, and V. V. Lobanenko. 2000. Functional association of CTCF with the insulator upstream of the H19 gene is parent of origin-specific and methylation-sensitive. *Curr Biol* **10**:853-856.
- Kapranov, P., S. E. Cawley, J. Drenkow, S. Bekiranov, R. L. Strausberg, S. P. Fodor, and T. R. Gingeras. 2002. Large-scale transcriptional activity in chromosomes 21 and 22. *Science* **296**:916-919.
- Karaman, M. W., M. L. Houck, L. G. Chemnick, S. Nagpal, D. Chawannakul, D. Sudano, B. L. Pike, V. V. Ho, O. A. Ryder, and J. G. Hacia. 2003. Comparative analysis of gene-expression patterns in human and African great ape cultured fibroblasts. *Genome Res* **13**:1619-1630.
- Kawaji, H., T. Kasukawa, S. Fukuda, S. Katayama, C. Kai, J. Kawai, P. Carninci, and Y. Hayashizaki. 2006. CAGE Basic/Analysis Databases: the CAGE resource for comprehensive promoter analysis. *Nucleic Acids Res* **34**:D632-636.
- Kel, A. E., E. Gossling, I. Reuter, E. Cheremushkin, O. V. Kel-Margoulis, and E. Wingender. 2003. MATCH: A tool for searching transcription factor binding sites in DNA sequences. *Nucleic Acids Res* **31**:3576-3579.
- Khaitovich, P., I. Hellmann, W. Enard, K. Nowick, M. Leinweber, H. Franz, G. Weiss, M. Lachmann, and S. Paabo. 2005. Parallel patterns of evolution in the genomes and transcriptomes of humans and chimpanzees. *Science* **309**:1850-1854.
- Khaitovich, P., S. Paabo, and G. Weiss. 2005. Toward a neutral evolutionary model of gene expression. *Genetics* **170**:929-939.
- Khambata-Ford, S., Y. Liu, C. Gleason, M. Dickson, R. B. Altman, S. Batzoglou, and R. M. Myers. 2003. Identification of promoter regions in the human genome by using a retroviral plasmid library-based functional reporter gene assay. *Genome Res* **13**:1765-1774.
- Khoury, G., and P. Gruss. 1983. Enhancer elements. *Cell* **33**:313-314.
- Kim, J., S. Klooster, and D. J. Shapiro. 1995. Intrinsically bent DNA in a eukaryotic transcription factor recognition sequence potentiates transcription activation. *J Biol Chem* **270**:1282-1288.
- Kim, T. H., L. O. Barrera, C. Qu, S. Van Calcar, N. D. Trinklein, S. J. Cooper, R. M. Luna, C. K. Glass, M. G. Rosenfeld, R. M. Myers, and B. Ren. 2005a. Direct isolation and identification of promoters in the human genome. *Genome Res* **15**:830-839.
- Kim, T. H., L. O. Barrera, M. Zheng, C. Qu, M. A. Singer, T. A. Richmond, Y. Wu, R. D. Green, and B. Ren. 2005b. A high-resolution map of active promoters in the human genome. *Nature* **436**:876-880.
- King, D. C., J. Taylor, L. Elnitski, F. Chiaromonte, W. Miller, and R. C. Hardison. 2005. Evaluation of regulatory potential and conservation scores for detecting cis-regulatory modules in aligned mammalian genome sequences. *Genome Res* **15**:1051-1060.
- King, M. C., and A. C. Wilson. 1975. Evolution at two levels in humans and chimpanzees. *Science* **188**:107-116.

- Kleinjan, D. A., A. Seawright, A. Schedl, R. A. Quinlan, S. Danes, and V. van Heyningen. 2001. Aniridia-associated translocations, DNase hypersensitivity, sequence comparison and transgenic analysis redefine the functional domain of PAX6. *Hum Mol Genet* **10**:2049-2059.
- Knight, J. C. 2005. Regulatory polymorphisms underlying complex disease traits. *J Mol Med* **83**:97-109.
- Knight, J. C., B. J. Keating, K. A. Rockett, and D. P. Kwiatkowski. 2003. In vivo characterization of regulatory polymorphisms by allele-specific quantification of RNA polymerase loading. *Nat Genet* **33**:469-475.
- Knudsen, S. 1999. Promoter2.0: for the recognition of PolIII promoter sequences. *Bioinformatics* **15**:356-361.
- Kolbe, D., J. Taylor, L. Elnitski, P. Eswara, J. Li, W. Miller, R. Hardison, and F. Chiaromonte. 2004. Regulatory potential scores from genome-wide three-way alignments of human, mouse, and rat. *Genome Res* **14**:700-707.
- Kong, S., D. Bohl, C. Li, and D. Tuan. 1997. Transcription of the HS2 enhancer toward a cis-linked gene is independent of the orientation, position, and distance of the enhancer relative to the gene. *Mol Cell Biol* **17**:3955-3965.
- Kumar, A., Y. Li, S. Patil, and S. Jain. 2005. A haplotype of the angiotensinogen gene is associated with hypertension in african americans. *Clin Exp Pharmacol Physiol* **32**:495-502.
- Kummerfeld, S. K., and S. A. Teichmann. 2006. DBD: a transcription factor prediction database. *Nucleic Acids Res* **34**:D74-81.
- Kutach, A. K., and J. T. Kadonaga. 2000. The downstream promoter element DPE appears to be as widely used as the TATA box in Drosophila core promoters. *Mol Cell Biol* **20**:4754-4764.
- Kuzmichev, A., and D. Reinberg. 2001. Role of histone deacetylase complexes in the regulation of chromatin metabolism. *Curr Top Microbiol Immunol* **254**:35-58.
- Lagrange, T., A. N. Kapanidis, H. Tang, D. Reinberg, and R. H. Ebright. 1998. New core promoter element in RNA polymerase II-dependent transcription: sequence-specific DNA binding by transcription factor IIB. *Genes Dev* **12**:34-44.
- Lagrange, T., T. K. Kim, G. Orphanides, Y. W. Ebright, R. H. Ebright, and D. Reinberg. 1996. High-resolution mapping of nucleoprotein complexes by site-specific protein-DNA photocrosslinking: organization of the human TBP-TFIIA-TFIIB-DNA quaternary complex. *Proc Natl Acad Sci U S A* **93**:10620-10625.
- Landy, A. 1989. Dynamic, structural, and regulatory aspects of lambda site-specific recombination. *Annu Rev Biochem* **58**:913-949.
- Lee, T. I., N. J. Rinaldi, F. Robert, D. T. Odom, Z. Bar-Joseph, G. K. Gerber, N. M. Hannett, C. T. Harbison, C. M. Thompson, I. Simon, J. Zeitlinger, E. G. Jennings, H. L. Murray, D. B. Gordon, B. Ren, J. J. Wyrick, J. B. Tagne, T. L. Volkert, E. Fraenkel, D. K. Gifford, and R. A. Young. 2002. Transcriptional regulatory networks in *Saccharomyces cerevisiae*. *Science* **298**:799-804.
- Legraverend, C., P. Antonson, P. Flodby, and K. G. Xanthopoulos. 1993. High level activity of the mouse CCAAT/enhancer binding protein (C/EBP alpha) gene promoter involves autoregulation and several ubiquitous transcription factors. *Nucleic Acids Res* **21**:1735-1742.
- Lemon, B., and R. Tjian. 2000. Orchestrated response: a symphony of transcription factors for gene control. *Genes Dev* **14**:2551-2569.

- Lerman, D. N., P. Michalak, A. B. Helin, B. R. Bettencourt, and M. E. Feder. 2003. Modification of heat-shock gene expression in *Drosophila melanogaster* populations via transposable elements. *Mol Biol Evol* **20**:135-144.
- Lettice, L. A., T. Horikoshi, S. J. Heaney, M. J. van Baren, H. C. van der Linde, G. J. Breedveld, M. Joosse, N. Akarsu, B. A. Oostra, N. Endo, M. Shibata, M. Suzuki, E. Takahashi, T. Shinka, Y. Nakahori, D. Ayusawa, K. Nakabayashi, S. W. Scherer, P. Heutink, R. E. Hill, and S. Noji. 2002. Disruption of a long-range cis-acting regulator for *Shh* causes preaxial polydactyly. *Proc Natl Acad Sci U S A* **99**:7548-7553.
- Lewin, B. 2003. *Genes VIII*. Prentice Hall.
- Li, Q., S. Harju, and K. R. Peterson. 1999. Locus control regions: coming of age at a decade plus. *Trends Genet* **15**:403-408.
- Li, S., F. C. MacLaughlin, J. G. Fewell, M. Gondo, J. Wang, F. Nicol, D. A. Dean, and L. C. Smith. 2001. Muscle-specific enhancement of gene expression by incorporation of SV40 enhancer in the expression plasmid. *Gene Ther* **8**:494-497.
- Li, Y., P. M. Flanagan, H. Tschochner, and R. D. Kornberg. 1994. RNA polymerase II initiation factor interactions and transcription start site selection. *Science* **263**:805-807.
- Li, Y., S. Jain, S. Patil, and A. Kumar. 2006. A haplotype of angiotensinogen gene that is associated with essential hypertension increases its promoter activity in adipocytes. *Vascul Pharmacol* **44**:29-33.
- Li, Z., S. Van Calcar, C. Qu, W. K. Cavenee, M. Q. Zhang, and B. Ren. 2003. A global transcriptional regulatory role for c-Myc in Burkitt's lymphoma cells. *Proc Natl Acad Sci U S A* **100**:8164-8169.
- Lim, C. Y., B. Santoso, T. Boulay, E. Dong, U. Ohler, and J. T. Kadonaga. 2004. The MTE, a new core promoter element for transcription by RNA polymerase II. *Genes Dev* **18**:1606-1617.
- Lin, C. H., S. Y. Hsieh, I. S. Sheen, W. C. Lee, T. C. Chen, W. C. Shyu, and Y. F. Liaw. 2001. Genome-wide hypomethylation in hepatocellular carcinogenesis. *Cancer Res* **61**:4238-4243.
- Lin, S., D. J. Cutler, M. E. Zwick, and A. Chakravarti. 2002. Haplotype inference in random population samples. *Am J Hum Genet* **71**:1129-1137.
- Linnell, J., R. Mott, S. Field, D. P. Kwiatkowski, J. Ragoussis, and I. A. Udalova. 2004. Quantitative high-throughput analysis of transcription factor binding specificities. *Nucleic Acids Res* **32**:e44.
- Lo, H. S., Z. Wang, Y. Hu, H. H. Yang, S. Gere, K. H. Buetow, and M. P. Lee. 2003. Allelic variation in gene expression is common in the human genome. *Genome Res* **13**:1855-1862.
- Ludwig, M. Z., C. Bergman, N. H. Patel, and M. Kreitman. 2000. Evidence for stabilizing selection in a eukaryotic enhancer element. *Nature* **403**:564-567.
- Martone, R., G. Euskirchen, P. Bertone, S. Hartman, T. E. Royce, N. M. Luscombe, J. L. Rinn, F. K. Nelson, P. Miller, M. Gerstein, S. Weissman, and M. Snyder. 2003. Distribution of NF-kappaB-binding sites across human chromosome 22. *Proc Natl Acad Sci U S A* **100**:12247-12252.
- Mastrangelo, I. A., A. J. Courey, J. S. Wall, S. P. Jackson, and P. V. Hough. 1991. DNA looping and Sp1 multimer links: a mechanism for transcriptional synergism and enhancement. *Proc Natl Acad Sci U S A* **88**:5670-5674.

- McGuire, W., A. V. Hill, C. E. Allsopp, B. M. Greenwood, and D. Kwiatkowski. 1994. Variation in the TNF-alpha promoter region associated with susceptibility to cerebral malaria. *Nature* **371**:508-510.
- McKusick, V. A. 1998. Mendelian Inheritance in Man. A Catalog of Human Genes and Genetic Disorders. Johns Hopkins University Press, Baltimore.
- Monks, S. A., A. Leonardson, H. Zhu, P. Cundiff, P. Pietrusiak, S. Edwards, J. W. Phillips, A. Sachs, and E. E. Schadt. 2004. Genetic inheritance of gene expression in human cell lines. *Am J Hum Genet* **75**:1094-1105.
- Montgomery, S. B., O. L. Griffith, M. C. Sleumer, C. M. Bergman, M. Bilenky, E. D. Pleasance, Y. Prychyna, X. Zhang, and S. J. Jones. 2006. ORegAnno: an open access database and curation system for literature-derived promoters, transcription factor binding sites and regulatory variation. *Bioinformatics* **22**:637-640.
- Morison, I. M., J. P. Ramsay, and H. G. Spencer. 2005. A census of mammalian imprinting. *Trends Genet* **21**:457-465.
- Morley, M., C. M. Molony, T. M. Weber, J. L. Devlin, K. G. Ewens, R. S. Spielman, and V. G. Cheung. 2004. Genetic analysis of genome-wide variation in human gene expression. *Nature* **430**:743-747.
- Mottagui-Tabar, S., M. A. Faghihi, Y. Mizuno, P. G. Engstrom, B. Lenhard, W. W. Wasserman, and C. Wahlestedt. 2005. Identification of functional SNPs in the 5-prime flanking sequences of human genes. *BMC Genomics* **6**:18.
- Nagaich, A. K., E. Appella, and R. E. Harrington. 1997. DNA bending is essential for the site-specific recognition of DNA response elements by the DNA binding domain of the tumor suppressor protein p53. *J Biol Chem* **272**:14842-14849.
- Nakashima, K., T. Hirota, K. Obara, M. Shimizu, S. Doi, K. Fujita, T. Shirakawa, T. Enomoto, S. Yoshihara, M. Ebisawa, K. Matsumoto, H. Saito, Y. Suzuki, Y. Nakamura, and M. Tamari. 2006. A functional polymorphism in MMP-9 is associated with childhood atopic asthma. *Biochem Biophys Res Commun* **344**:300-307.
- Nightingale, K. P., L. P. O'Neill, and B. M. Turner. 2006. Histone modifications: signalling receptors and potential elements of a heritable epigenetic code. *Curr Opin Genet Dev* **16**:125-136.
- Nikolov, D. B., H. Chen, E. D. Halay, A. A. Usheva, K. Hisatake, D. K. Lee, R. G. Roeder, and S. K. Burley. 1995. Crystal structure of a TFIIB-TBP-TATA-element ternary complex. *Nature* **377**:119-128.
- Nishikawa, J., M. Amano, Y. Fukue, S. Tanaka, H. Kishi, Y. Hirota, K. Yoda, and T. Ohyama. 2003. Left-handedly curved DNA regulates accessibility to cis-DNA elements in chromatin. *Nucleic Acids Res* **31**:6651-6662.
- Niu, T., Z. S. Qin, X. Xu, and J. S. Liu. 2002. Bayesian haplotype inference for multiple linked single-nucleotide polymorphisms. *Am J Hum Genet* **70**:157-169.
- O'Shea-Greenfield, A., and S. T. Smale. 1992. Roles of TATA and initiator elements in determining the start site location and direction of RNA polymerase II transcription. *J Biol Chem* **267**:6450.
- Odom, D. T., N. Zizlsperger, D. B. Gordon, G. W. Bell, N. J. Rinaldi, H. L. Murray, T. L. Volkert, J. Schreiber, P. A. Rolfe, D. K. Gifford, E. Fraenkel, G. I. Bell, and R. A. Young. 2004. Control of pancreas and liver gene expression by HNF transcription factors. *Science* **303**:1378-1381.
- Ohler, U., G. C. Liao, H. Niemann, and G. M. Rubin. 2002. Computational analysis of core promoters in the Drosophila genome. *Genome Biol* **3**:RESEARCH0087.

- Park, J. Y., J. M. Park, J. S. Jang, J. E. Choi, K. M. Kim, S. I. Cha, C. H. Kim, Y. M. Kang, W. K. Lee, S. Kam, R. W. Park, I. S. Kim, J. T. Lee, and T. H. Jung. 2006. Caspase 9 promoter polymorphisms and risk of primary lung cancer. *Hum Mol Genet* **15**:1963-1971.
- Parvin, J. D., R. J. McCormick, P. A. Sharp, and D. E. Fisher. 1995. Pre-bending of a promoter sequence enhances affinity for the TATA-binding factor. *Nature* **373**:724-727.
- Pastinen, T., B. Ge, S. Gurd, T. Gaudin, C. Dore, M. Lemire, P. Lepage, E. Harmsen, and T. J. Hudson. 2005. Mapping common regulatory variants to human haplotypes. *Hum Mol Genet* **14**:3963-3971.
- Pastinen, T., B. Ge, and T. J. Hudson. 2006. Influence of human genome polymorphism on gene expression. *Hum Mol Genet* **15 Suppl 1**:R9-R16.
- Pastinen, T., and T. J. Hudson. 2004. Cis-acting regulatory variation in the human genome. *Science* **306**:647-650.
- Pastinen, T., R. Sladek, S. Gurd, A. Sammak, B. Ge, P. Lepage, K. Lavergne, A. Villeneuve, T. Gaudin, H. Brandstrom, A. Beck, A. Verner, J. Kingsley, E. Harmsen, D. Labuda, K. Morgan, M. C. Vohl, A. K. Naumova, D. Sinnett, and T. J. Hudson. 2004. A survey of genetic and epigenetic variation affecting human gene expression. *Physiol Genomics* **16**:184-193.
- Pazin, M. J., and J. T. Kadonaga. 1997. What's up and down with histone deacetylation and transcription? *Cell* **89**:325-328.
- Ponger, L., and D. Mouchiroud. 2002. CpGProD: identifying CpG islands associated with transcription start sites in large genomic mammalian sequences. *Bioinformatics* **18**:631-633.
- Pritchard, J. K., and M. Przeworski. 2001. Linkage disequilibrium in humans: models and data. *Am J Hum Genet* **69**:1-14.
- Przeworski, M., R. R. Hudson, and A. Di Rienzo. 2000. Adjusting the focus on human variation. *Trends Genet* **16**:296-302.
- Ptashne, M. 1992. *A Genetic Switch: Phage (Lambda) and Higher Organisms*. Cell Press, Cambridge, MA.
- Ragoczy, T., A. Telling, T. Sawado, M. Groudine, and S. T. Kosak. 2003. A genetic analysis of chromosome territory looping: diverse roles for distal regulatory elements. *Chromosome Res* **11**:513-525.
- Rastegar, M., F. P. Lemaigre, and G. G. Rousseau. 2000. Control of gene expression by growth hormone in liver: key role of a network of transcription factors. *Mol Cell Endocrinol* **164**:1-4.
- Reese, M. G. 2001. Application of a time-delay neural network to promoter annotation in the *Drosophila melanogaster* genome. *Comput Chem* **26**:51-56.
- Reik, W., and J. Walter. 2001. Genomic imprinting: parental influence on the genome. *Nat Rev Genet* **2**:21-32.
- Reiter, W. D., U. Hudepohl, and W. Zillig. 1990. Mutational analysis of an archaeobacterial promoter: essential role of a TATA box for transcription efficiency and start-site selection in vitro. *Proc Natl Acad Sci U S A* **87**:9509-9513.
- Ren, B., F. Robert, J. J. Wyrick, O. Aparicio, E. G. Jennings, I. Simon, J. Zeitlinger, J. Schreiber, N. Hannett, E. Kanin, T. L. Volkert, C. J. Wilson, S. P. Bell, and R. A. Young. 2000. Genome-wide location and function of DNA binding proteins. *Science* **290**:2306-2309.

- Rezai-Zadeh, N., X. Zhang, F. Namour, G. Fejer, Y. D. Wen, Y. L. Yao, I. Gyory, K. Wright, and E. Seto. 2003. Targeted recruitment of a histone H4-specific methyltransferase by the transcription factor YY1. *Genes Dev* **17**:1019-1029.
- Rhodes, S. J., R. Chen, G. E. DiMattia, K. M. Scully, K. A. Kalla, S. C. Lin, V. C. Yu, and M. G. Rosenfeld. 1993. A tissue-specific enhancer confers Pit-1-dependent morphogen inducibility and autoregulation on the pit-1 gene. *Genes Dev* **7**:913-932.
- Rice, J. C., S. D. Briggs, B. Ueberheide, C. M. Barber, J. Shabanowitz, D. F. Hunt, Y. Shinkai, and C. D. Allis. 2003. Histone methyltransferases direct different degrees of methylation to define distinct chromatin domains. *Mol Cell* **12**:1591-1598.
- Rifkin, S. A., J. Kim, and K. P. White. 2003. Evolution of gene expression in the *Drosophila melanogaster* subgroup. *Nat Genet* **33**:138-144.
- Rinn, J. L., G. Euskirchen, P. Bertone, R. Martone, N. M. Luscombe, S. Hartman, P. M. Harrison, F. K. Nelson, P. Miller, M. Gerstein, S. Weissman, and M. Snyder. 2003. The transcriptional activity of human Chromosome 22. *Genes Dev* **17**:529-540.
- Robertson, G., M. Bilenky, K. Lin, A. He, W. Yuen, M. Dagpinar, R. Varhol, K. Teague, O. L. Griffith, X. Zhang, Y. Pan, M. Hassel, M. C. Sleumer, W. Pan, E. D. Pleasance, M. Chuang, H. Hao, Y. Y. Li, N. Robertson, C. Fjell, B. Li, S. B. Montgomery, T. Astakhova, J. Zhou, J. Sander, A. S. Siddiqui, and S. J. Jones. 2006. cisRED: a database system for genome-scale computational discovery of regulatory elements. *Nucleic Acids Res* **34**:D68-73.
- Rockman, M. V., and G. A. Wray. 2002. Abundant raw material for cis-regulatory evolution in humans. *Mol Biol Evol* **19**:1991-2004.
- Ross, D. T., U. Scherf, M. B. Eisen, C. M. Perou, C. Rees, P. Spellman, V. Iyer, S. S. Jeffrey, M. Van de Rijn, M. Waltham, A. Pergamenschikov, J. C. Lee, D. Lashkari, D. Shalon, T. G. Myers, J. N. Weinstein, D. Botstein, and P. O. Brown. 2000. Systematic variation in gene expression patterns in human cancer cell lines. *Nat Genet* **24**:227-235.
- Roth, F. P., J. D. Hughes, P. W. Estep, and G. M. Church. 1998. Finding DNA regulatory motifs within unaligned noncoding sequences clustered by whole-genome mRNA quantitation. *Nat Biotechnol* **16**:939-945.
- Roth, S. Y., J. M. Denu, and C. D. Allis. 2001. Histone acetyltransferases. *Annu Rev Biochem* **70**:81-120.
- Rothenburg, S., F. Koch-Nolte, A. Rich, and F. Haag. 2001a. A polymorphic dinucleotide repeat in the rat nucleolin gene forms Z-DNA and inhibits promoter activity. *Proc Natl Acad Sci U S A* **98**:8985-8990.
- Rothenburg, S., F. Koch-Nolte, H. G. Thiele, and F. Haag. 2001b. DNA methylation contributes to tissue- and allele-specific expression of the T-cell differentiation marker RT6. *Immunogenetics* **52**:231-241.
- Rozen, S., and H. Skaletsky. 2000. Primer3 on the WWW for general users and for biologist programmers. *Methods Mol Biol* **132**:365-386.
- Sachidanandam, R., D. Weissman, S. C. Schmidt, J. M. Kakol, L. D. Stein, G. Marth, S. Sherry, J. C. Mullikin, B. J. Mortimore, D. L. Willey, S. E. Hunt, C. G. Cole, P. C. Coggill, C. M. Rice, Z. Ning, J. Rogers, D. R. Bentley, P. Y. Kwok, E. R. Mardis, R. T. Yeh, B. Schultz, L. Cook, R. Davenport, M. Dante, L. Fulton, L. Hillier, R. H. Waterston, J. D. McPherson, B. Gilman, S. Schaffner, W. J. Van Etten, D. Reich, J. Higgins, M. J. Daly, B. Blumenstiel, J. Baldwin, N. Stange-Thomann, M. C. Zody, L. Linton, E. S. Lander, and D.

- Altshuler. 2001. A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms. *Nature* **409**:928-933.
- Saito, T., F. Guan, D. F. Papolos, N. Rajouria, C. S. Fann, and H. M. Lachman. 2001. Polymorphism in SNAP29 gene promoter region associated with schizophrenia. *Mol Psychiatry* **6**:193-201.
- Sandelin, A., W. Alkema, P. Engstrom, W. W. Wasserman, and B. Lenhard. 2004. JASPAR: an open-access database for eukaryotic transcription factor binding profiles. *Nucleic Acids Res* **32**:D91-94.
- Sawado, T., J. Halow, M. A. Bender, and M. Groudine. 2003. The beta -globin locus control region (LCR) functions primarily by enhancing the transition from transcription initiation to elongation. *Genes Dev* **17**:1009-1018.
- Schubeler, D., M. Groudine, and M. A. Bender. 2001. The murine beta-globin locus control region regulates the rate of transcription but not the hyperacetylation of histones at the active genes. *Proc Natl Acad Sci U S A* **98**:11432-11437.
- Schubeler, D., D. M. MacAlpine, D. Scalzo, C. Wirbelauer, C. Kooperberg, F. van Leeuwen, D. E. Gottschling, L. P. O'Neill, B. M. Turner, J. Delrow, S. P. Bell, and M. Groudine. 2004. The histone modification pattern of active genes revealed through genome-wide chromatin analysis of a higher eukaryote. *Genes Dev* **18**:1263-1271.
- Seenisamy, J., E. M. Rezler, T. J. Powell, D. Tye, V. Gokhale, C. S. Joshi, A. Siddiqui-Jain, and L. H. Hurley. 2004. The dynamic character of the G-quadruplex element in the c-MYC promoter and modification by TMPyP4. *J Am Chem Soc* **126**:8702-8709.
- Segal, J. A., J. L. Barnett, and D. L. Crawford. 1999. Functional analyses of natural variation in Sp1 binding sites of a TATA-less promoter. *J Mol Evol* **49**:736-749.
- Shaw, M. A., I. J. Donaldson, A. Collins, C. S. Peacock, Z. Lins-Lainson, J. J. Shaw, F. Ramos, F. Silveira, and J. M. Blackwell. 2001. Association and linkage of leprosy phenotypes with HLA class II and tumour necrosis factor genes. *Genes Immun* **2**:196-204.
- Shedden, K., W. Chen, R. Kuick, D. Ghosh, J. Macdonald, K. R. Cho, T. J. Giordano, S. B. Gruber, E. R. Fearon, J. M. Taylor, and S. Hanash. 2005. Comparison of seven methods for producing Affymetrix expression scores based on False Discovery Rates in disease profiling data. *BMC Bioinformatics* **6**:26.
- Shin, H. D., C. Winkler, J. C. Stephens, J. Bream, H. Young, J. J. Goedert, T. R. O'Brien, D. Vlahov, S. Buchbinder, J. Giorgi, C. Rinaldo, S. Donfield, A. Willoughby, S. J. O'Brien, and M. W. Smith. 2000. Genetic restriction of HIV-1 pathogenesis to AIDS by promoter alleles of IL10. *Proc Natl Acad Sci U S A* **97**:14467-14472.
- Shiraki, T., S. Kondo, S. Katayama, K. Waki, T. Kasukawa, H. Kawaji, R. Kodzius, A. Watahiki, M. Nakamura, T. Arakawa, S. Fukuda, D. Sasaki, A. Podhajska, M. Harbers, J. Kawai, P. Carninci, and Y. Hayashizaki. 2003. Cap analysis gene expression for high-throughput analysis of transcriptional starting point and identification of promoter usage. *Proc Natl Acad Sci U S A* **100**:15776-15781.
- Siepel, A., G. Bejerano, J. S. Pedersen, A. S. Hinrichs, M. Hou, K. Rosenbloom, H. Clawson, J. Spieth, L. W. Hillier, S. Richards, G. M. Weinstock, R. K. Wilson, R. A. Gibbs, W. J. Kent, W. Miller, and D. Haussler. 2005. Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. *Genome Res* **15**:1034-1050.

- Smale, S. T. 1997. Transcription initiation from TATA-less promoters within eukaryotic protein-coding genes. *Biochim Biophys Acta* **1351**:73-88.
- Smale, S. T., and D. Baltimore. 1989. The "initiator" as a transcription control element. *Cell* **57**:103-113.
- Snoussi, K., W. Mahfoudh, N. Bouaouina, S. B. Ahmed, A. N. Helal, and L. Chouchane. 2006. Genetic variation in IL-8 associated with increased risk and poor prognosis of breast carcinoma. *Hum Immunol* **67**:13-21.
- Spellman, P. T., G. Sherlock, M. Q. Zhang, V. R. Iyer, K. Anders, M. B. Eisen, P. O. Brown, D. Botstein, and B. Futcher. 1998. Comprehensive identification of cell cycle-regulated genes of the yeast *Saccharomyces cerevisiae* by microarray hybridization. *Mol Biol Cell* **9**:3273-3297.
- Spiecker, M., H. Darius, T. Hankeln, M. Soufi, A. M. Sattler, J. R. Schaefer, K. Node, J. Borgel, A. Mugge, K. Lindpaintner, A. Huesing, B. Maisch, D. C. Zeldin, and J. K. Liao. 2004. Risk of coronary artery disease associated with polymorphism of the cytochrome P450 epoxygenase CYP2J2. *Circulation* **110**:2132-2136.
- Spitz, F., F. Gonzalez, and D. Duboule. 2003. A global control region defines a chromosomal regulatory landscape containing the HoxD cluster. *Cell* **113**:405-417.
- Sprecher, E., A. Ishida-Yamamoto, M. Mizrahi-Koren, D. Rapaport, D. Goldsher, M. Indelman, O. Topaz, I. Chefetz, H. Keren, J. O'Brien T, D. Bercovich, S. Shalev, D. Geiger, R. Bergman, M. Horowitz, and H. Mandel. 2005. A mutation in SNAP29, coding for a SNARE protein involved in intracellular trafficking, causes a novel neurocutaneous syndrome characterized by cerebral dysgenesis, neuropathy, ichthyosis, and palmoplantar keratoderma. *Am J Hum Genet* **77**:242-251.
- Stanford, W. L., J. B. Cohn, and S. P. Cordes. 2001. Gene-trap mutagenesis: past, present and beyond. *Nat Rev Genet* **2**:756-768.
- Stephens, M., and P. Scheet. 2005. Accounting for decay of linkage disequilibrium in haplotype inference and missing-data imputation. *Am J Hum Genet* **76**:449-462.
- Stephens, M., N. J. Smith, and P. Donnelly. 2001. A new statistical method for haplotype reconstruction from population data. *Am J Hum Genet* **68**:978-989.
- Sterner, D. E., and S. L. Berger. 2000. Acetylation of histones and transcription-related factors. *Microbiol Mol Biol Rev* **64**:435-459.
- Stranger, B. E., M. S. Forrest, A. G. Clark, M. J. Minichiello, S. Deutsch, R. Lyle, S. Hunt, B. Kahl, S. E. Antonarakis, S. Tavaré, P. Deloukas, and E. T. Dermitzakis. 2005. Genome-Wide Associations of Gene Expression Variation in Humans. *PLoS Genet* **1**:e78.
- Strathdee, G., A. Sim, and R. Brown. 2004. Control of gene expression by CpG island methylation in normal cells. *Biochem Soc Trans* **32**:913-915.
- Strathdee, G., A. Sim, R. Soutar, T. L. Holyoake, and R. Brown. 2006. HOXA5 is targeted by cell type specific CpG island methylation in normal cells and during the development of acute myeloid leukaemia. *Carcinogenesis*.
- Suzuki, M., and N. Yagi. 1995. Stereochemical basis of DNA bending by transcription factors. *Nucleic Acids Res* **23**:2083-2091.
- Suzuki, Y., T. Tsunoda, J. Sese, H. Taira, J. Mizushima-Sugano, H. Hata, T. Ota, T. Isogai, T. Tanaka, Y. Nakamura, A. Suyama, Y. Sakaki, S. Morishita, K. Okubo, and S. Sugano. 2001. Identification and characterization of the

- potential promoter regions of 1031 kinds of human genes. *Genome Res* **11**:677-684.
- Suzuki, Y., R. Yamashita, M. Shirota, Y. Sakakibara, J. Chiba, J. Mizushima-Sugano, K. Nakai, and S. Sugano. 2004. Sequence comparison of human and mouse genes reveals a homologous block structure in the promoter regions. *Genome Res* **14**:1711-1718.
- Talkington, C. A., and P. Leder. 1982. Rescuing the in vitro function of a globin pseudogene promoter. *Nature* **298**:192-195.
- Thijs, G., K. Marchal, M. Lescot, S. Rombauts, B. De Moor, P. Rouze, and Y. Moreau. 2002. A Gibbs sampling method to detect overrepresented motifs in the upstream regions of coexpressed genes. *J Comput Biol* **9**:447-464.
- Timchenko, N., D. R. Wilson, L. R. Taylor, S. Abdelsayed, M. Wilde, M. Sawadogo, and G. J. Darlington. 1995. Autoregulation of the human C/EBP alpha gene by stimulation of upstream stimulatory factor binding. *Mol Cell Biol* **15**:1192-1202.
- Ting, J. P., and J. Trowsdale. 2002. Genetic control of MHC class II expression. *Cell* **109 Suppl**:S21-33.
- Tolhuis, B., R. J. Palstra, E. Splinter, F. Grosveld, and W. de Laat. 2002. Looping and interaction between hypersensitive sites in the active beta-globin locus. *Mol Cell* **10**:1453-1465.
- Tournamille, C., Y. Colin, J. P. Cartron, and C. Le Van Kim. 1995. Disruption of a GATA motif in the Duffy gene promoter abolishes erythroid gene expression in Duffy-negative individuals. *Nat Genet* **10**:224-228.
- Trinklein, N. D., S. F. Aldred, S. J. Hartman, D. I. Schroeder, R. P. Otilar, and R. M. Myers. 2004. An abundance of bidirectional promoters in the human genome. *Genome Res* **14**:62-66.
- Trinklein, N. D., S. J. Aldred, A. J. Saldanha, and R. M. Myers. 2003. Identification and functional analysis of human transcriptional promoters. *Genome Res* **13**:308-312.
- Umbricht, C. B., E. Evron, E. Gabrielson, A. Ferguson, J. Marks, and S. Sukumar. 2001. Hypermethylation of 14-3-3 sigma (stratifin) is an early event in breast cancer. *Oncogene* **20**:3348-3353.
- Wang, R. L., A. Stec, J. Hey, L. Lukens, and J. Doebley. 1999. The limits of selection during maize domestication. *Nature* **398**:236-239.
- Warren, P., J. Bienkowska, P. G. V. Martini, J. Jackson, and D. M. Taylor. 2006. PANP - a New Method of Gene Detection on Oligonucleotide Expression Arrays. *Research in Computational Molecular Biology, Venice, Italy*.
- Wasserman, W. W., and A. Sandelin. 2004. Applied bioinformatics for the identification of regulatory elements. *Nat Rev Genet* **5**:276-287.
- Wasylyk, B., R. Derbyshire, A. Guy, D. Molko, A. Roget, R. Teoule, and P. Chambon. 1980. Specific in vitro transcription of conalbumin gene is drastically decreased by single-point mutation in T-A-T-A box homology sequence. *Proc Natl Acad Sci U S A* **77**:7024-7028.
- Weisberg, R., and A. Landy. 1983. Site-Specific Recombination in Phage Lambda. Pp. 211-250 *in* R. Weisberg, ed. *Lambda II*. Cold Spring Harbor Press, Cold Spring Harbor, NY.
- West, A. G., and P. Fraser. 2005. Remote control of gene transcription. *Hum Mol Genet* **14 Spec No 1**:R101-111.
- West, A. G., M. Gaszner, and G. Felsenfeld. 2002. Insulators: many functions, many mechanisms. *Genes Dev* **16**:271-288.

- Whitehead, A. S., and R. Sackstein. 1985. Molecular biology of the human and mouse MHC class III genes: phylogenetic conservation, genetics and regulation of expression. *Immunol Rev* **87**:185-208.
- Wingender, E., P. Dietze, H. Karas, and R. Knuppel. 1996. TRANSFAC: a database on transcription factors and their DNA binding sites. *Nucleic Acids Res* **24**:238-241.
- Wittkopp, P. J., B. K. Haerum, and A. G. Clark. 2004. Evolutionary changes in cis and trans gene regulation. *Nature* **430**:85-88.
- Wonodi, I., L. E. Hong, M. T. Avila, R. W. Buchanan, W. T. Carpenter, Jr., O. C. Stine, B. D. Mitchell, and G. K. Thaker. 2005. Association between polymorphism of the SNAP29 gene promoter region and schizophrenia. *Schizophr Res* **78**:339-341.
- Wray, G. A., M. W. Hahn, E. Abouheif, J. P. Balhoff, M. Pizer, M. V. Rockman, and L. A. Romano. 2003. The evolution of transcriptional regulation in eukaryotes. *Mol Biol Evol* **20**:1377-1419.
- Wright, S., E. deBoer, A. Rosenthal, R. A. Flavell, and F. Grosveld. 1984. DNA sequences required for regulated expression of beta-globin genes in murine erythroleukaemia cells. *Philos Trans R Soc Lond B Biol Sci* **307**:271-282.
- Wu, Z., R. Irizarry, R. Gentleman, F. Martinez-Murillo, and F. Spencer. 2004. A Model-Based Background Adjustment for Oligonucleotide Expression Arrays. *Journal of the American Statistical Association* **99**:909-917.
- Xie, X., J. Lu, E. J. Kulbokas, T. R. Golub, V. Mootha, K. Lindblad-Toh, E. S. Lander, and M. Kellis. 2005. Systematic discovery of regulatory motifs in human promoters and 3' UTRs by comparison of several mammals. *Nature* **434**:338-345.
- Xu, L. C., M. Thali, and W. Schaffner. 1991. Upstream box/TATA box order is the major determinant of the direction of transcription. *Nucleic Acids Res* **19**:6699-6704.
- Yamamoto, K. R. 1985. Steroid receptor regulated transcription of specific genes and gene networks. *Annu Rev Genet* **19**:209-252.
- Yan, H., Z. Dobbie, S. B. Gruber, S. Markowitz, K. Romans, F. M. Giardiello, K. W. Kinzler, and B. Vogelstein. 2002a. Small changes in expression affect predisposition to tumorigenesis. *Nat Genet* **30**:25-26.
- Yan, H., W. Yuan, V. E. Velculescu, B. Vogelstein, and K. W. Kinzler. 2002b. Allelic variation in human gene expression. *Science* **297**:1143.