

## Reference

- Aerts, S., P. Van Loo, et al. (2005). "TOUCAN 2: the all-inclusive open source workbench for regulatory sequence analysis." Nucleic Acids Res **33**(Web Server issue): W393-6.
- Allen, T. A., S. Von Kaenel, et al. (2004). "The SINE-encoded mouse B2 RNA represses mRNA transcription in response to heat shock." Nat Struct Mol Biol **11**(9): 816-21.
- Anderson, R. P. and J. R. Roth (1977). "Tandem genetic duplications in phage and bacteria." Annu Rev Microbiol **31**: 473-505.
- Arney, K. L. (2003). "H19 and Igf2--enhancing the confusion?" Trends Genet **19**(1): 17-23.
- Bafna, V. and S. Zhang (2004). "FastR: fast database search tool for non-coding RNA." Proc IEEE Comput Syst Bioinform Conf: 52-61.
- Bailey, T. L. and C. Elkan (1994). "Fitting a mixture model by expectation maximization to discover motifs in biopolymers." Proc Int Conf Intell Syst Mol Biol **2**: 28-36.
- Bailey, T. L. and W. S. Noble (2003). "Searching for statistically significant regulatory modules." Bioinformatics **19 Suppl 2**: ii16-25.
- Barash, Y., G. Elidan, et al. (2003). Modeling dependencies in protein-dna binding sites. In Proceedings of Seventh Annual International Conference on Computational Molecular Biology (RECOMB), ACM press, New York: 28-37.
- Bartel, D. P. (2004). "MicroRNAs: genomics, biogenesis, mechanism, and function." Cell **116**(2): 281-97.
- Baskerville, S. and D. P. Bartel (2005). "Microarray profiling of microRNAs reveals frequent coexpression with neighboring miRNAs and host genes." Rna **11**(3): 241-7.
- Beckmann, J. S., P. F. Johnson, et al. (1977). "Cloning of yeast transfer RNA genes in Escherichia coli." Science **196**(4286): 205-8.
- Bentwich, I., A. Avniel, et al. (2005). "Identification of hundreds of conserved and nonconserved human microRNAs." Nat Genet **37**(7): 766-70.
- Berg, O. G. and P. H. von Hippel (1987). "Selection of DNA binding sites by regulatory proteins. Statistical-mechanical theory and application to operators and promoters." J Mol Biol **193**(4): 723-50.
- Berger, S. L. (2002). "Histone modifications in transcriptional regulation." Curr Opin Genet Dev **12**(2): 142-8.
- Birney, E., D. Andrews, et al. (2004). "Ensembl 2004." Nucleic Acids Res **32 Database issue**: D468-70.
- Birney, E., D. Andrews, et al. (2006). "Ensembl 2006." Nucleic Acids Res **34**(Database issue): D556-61.
- Borer, P. N., B. Dengler, et al. (1974). "Stability of ribonucleic acid double-stranded helices." J Mol Biol **86**(4): 843-53.
- Brendel, V., G. H. Hamm, et al. (1986). "Terminators of transcription with RNA polymerase

- from Escherichia coli: what they look like and how to find them." J Biomol Struct Dyn **3**(4): 705-23.
- Brent, M. R. (2005). "Genome annotation past, present, and future: how to define an ORF at each locus." Genome Res **15**(12): 1777-86.
- Burge, C. and S. Karlin (1997). "Prediction of complete gene structures in human genomic DNA." J Mol Biol **268**(1): 78-94.
- Camier, S., R. E. Baker, et al. (1990). "On the flexible interaction of yeast factor tau with the bipartite promoter of tRNA genes." Nucleic Acids Res **18**(15): 4571-8.
- Cannon, R. E., G. J. Wu, et al. (1986). "Functions of and interactions between the A and B blocks in adenovirus type 2-specific VARNA1 gene." Proc Natl Acad Sci U S A **83**(5): 1285-9.
- Carninci, P., T. Kasukawa, et al. (2005). "The transcriptional landscape of the mammalian genome." Science **309**(5740): 1559-63.
- Cavarelli, J., B. Rees, et al. (1993). "Yeast tRNA(Asp) recognition by its cognate class II aminoacyl-tRNA synthetase." Nature **362**(6416): 181-4.
- Cech, T. R., A. J. Zaugg, et al. (1981). "In vitro splicing of the ribosomal RNA precursor of Tetrahymena: involvement of a guanosine nucleotide in the excision of the intervening sequence." Cell **27**(3 Pt 2): 487-96.
- Chang, Y. N., I. L. Pirtle, et al. (1986). "Nucleotide sequence and transcription of a human tRNA gene cluster with four genes." Gene **48**(1): 165-74.
- Chen, J. H., S. Y. Le, et al. (2000). "Prediction of common secondary structures of RNAs: a genetic algorithm approach." Nucleic Acids Res **28**(4): 991-9.
- Cheung, J., X. Estivill, et al. (2003). "Genome-wide detection of segmental duplications and potential assembly errors in the human genome sequence." Genome Biol **4**(4): R25.
- Chiang, D., A. K. Joshi, et al. (2006). "Grammatical representations of macromolecular structure." J Comput Biol **13**(5): 1077-100.
- Chiu, D. K. and T. Kolodziejczak (1991). "Inferring consensus structure from nucleic acid sequences." Comput Appl Biosci **7**(3): 347-52.
- Chomsky, D. (1959). "On certain formal properties of grammars." Inform. Cont. **2**: 137-176.
- Clamp, M., D. Andrews, et al. (2003). "Ensembl 2002: accommodating comparative genomics." Nucleic Acids Res **31**(1): 38-42.
- Clote, P., F. Ferre, et al. (2005). "Structural RNA has lower folding energy than random RNA of the same dinucleotide frequency." Rna **11**(5): 578-91.
- Coventry, A., D. J. Kleitman, et al. (2004). "MSARI: multiple sequence alignments for statistical detection of RNA secondary structure." Proc Natl Acad Sci U S A **101**(33): 12102-7.
- d'Aubenton Carafa, Y., E. Brody, et al. (1990). "Prediction of rho-independent Escherichia coli transcription terminators. A statistical analysis of their RNA stem-loop structures." J

- Mol Biol **216**(4): 835-58.
- de Hoon, M. J., Y. Makita, et al. (2005). "Prediction of Transcriptional Terminators in *Bacillus subtilis* and Related Species." PLoS Comput Biol **1**(3): e25.
- Dean, A. (2006). "On a chromosome far, far away: LCRs and gene expression." Trends Genet **22**(1): 38-45.
- DeFranco, D., O. Schmidt, et al. (1980). "Two control regions for eukaryotic tRNA gene transcription." Proc Natl Acad Sci U S A **77**(6): 3365-8.
- Dehal, P. and J. L. Boore (2005). "Two rounds of whole genome duplication in the ancestral vertebrate." PLoS Biol **3**(10): e314.
- di Bernardo, D., T. Down, et al. (2003). "ddbRNA: detection of conserved secondary structures in multiple alignments." Bioinformatics **19**(13): 1606-11.
- Dirks, R. M. and N. A. Pierce (2003). "A partition function algorithm for nucleic acid secondary structure including pseudoknots." J Comput Chem **24**(13): 1664-77.
- Dittmar, K. A., J. M. Goodenbour, et al. (2006). "Tissue-Specific Differences in Human Transfer RNA Expression." PLoS Genet **2**(12): e221.
- Domitrovich, A. M. and G. R. Kunkel (2003). "Multiple, dispersed human U6 small nuclear RNA genes with varied transcriptional efficiencies." Nucleic Acids Res **31**(9): 2344-52.
- Down, T., B. Leong, et al. (2006). "A machine learning strategy to identify candidate binding sites in human protein-coding sequence." BMC Bioinformatics **7**: 419.
- Down, T. A. (2002). Computational localization of promoters and transcription start sites in mammalian genomes. The Wellcome Trust Sanger Institute. Hinxton, Cambridge, Cambridge.
- Down, T. A. and T. J. Hubbard (2002). "Computational detection and location of transcription start sites in mammalian genomic DNA." Genome Res **12**(3): 458-61.
- Down, T. A. and T. J. Hubbard (2004). "What can we learn from noncoding regions of similarity between genomes?" BMC Bioinformatics **5**: 131.
- Durbin, R., S. R. Eddy, et al. (1998). Biological Sequence Analysis: Probabilistic Models of Proteins and Nucleic Acids. Cambridge, Cambridge University Press.
- Eddy, S. R. (2004). "How do RNA folding algorithms work?" Nat Biotechnol **22**(11): 1457-8.
- Eddy, S. R. and R. Durbin (1994). "RNA sequence analysis using covariance models." Nucleic Acids Res **22**(11): 2079-88.
- Ermolaeva, M. D., H. G. Khalak, et al. (2000). "Prediction of transcription terminators in bacterial genomes." J Mol Biol **301**(1): 27-33.
- Espinoza, C. A., T. A. Allen, et al. (2004). "B2 RNA binds directly to RNA polymerase II to repress transcript synthesis." Nat Struct Mol Biol **11**(9): 822-9.
- Farnham, P. J. and T. Platt (1981). "Rho-independent termination: dyad symmetry in DNA causes RNA polymerase to pause during transcription in vitro." Nucleic Acids Res **9**(3):

563-77.

- Fatica, A., M. Morlando, et al. (2000). "Yeast snoRNA accumulation relies on a cleavage-dependent/polyadenylation-independent 3'-processing apparatus." Embo J **19**(22): 6218-29.
- Fichant, G. A. and C. Burks (1991). "Identifying potential tRNA genes in genomic DNA sequences." J Mol Biol **220**(3): 659-71.
- Fickett, J. W. and A. G. Hatzigeorgiou (1997). "Eukaryotic promoter recognition." Genome Res **7**(9): 861-78.
- Fields, D. S. and R. R. Gutell (1996). "An analysis of large rRNA sequences folded by a thermodynamic method." Fold Des **1**(6): 419-30.
- Flamm, C., W. Fontana, et al. (2000). "RNA folding at elementary step resolution." Rna **6**(3): 325-38.
- Fournier, M. J., W. L. Miller, et al. (1974). "Clustering of tRNA cistrons in Escherichia coli DNA." Biochem Biophys Res Commun **60**(3): 1148-54.
- Frazer, K. A., L. Pachter, et al. (2004). "VISTA: computational tools for comparative genomics." Nucleic Acids Res **32**(Web Server issue): W273-9.
- Galli, G., H. Hofstetter, et al. (1981). "Two conserved sequence blocks within eukaryotic tRNA genes are major promoter elements." Nature **294**(5842): 626-31.
- Gardner, P. P. and R. Giegerich (2004). "A comprehensive comparison of comparative RNA structure prediction approaches." BMC Bioinformatics **5**(1): 140.
- Gautheret, D. and A. Lambert (2001). "Direct RNA motif definition and identification from multiple sequence alignments using secondary structure profiles." J Mol Biol **313**(5): 1003-11.
- Giege, R., M. Sissler, et al. (1998). "Universal rules and idiosyncratic features in tRNA identity." Nucleic Acids Res **26**(22): 5017-35.
- Giles, K. E., M. Caputi, et al. (2004). "Packaging and reverse transcription of snRNAs by retroviruses may generate pseudogenes." Rna **10**(2): 299-307.
- Gish, W. (1996-2004). "WU-BLAST." <http://blast.wustl.edu>.
- Gopalan, V., A. Vioque, et al. (2002). "RNase P: variations and uses." J Biol Chem **277**(9): 6759-62.
- Gorodkin, J., S. L. Stricklin, et al. (2001). "Discovering common stem-loop motifs in unaligned RNA sequences." Nucleic Acids Res **29**(10): 2135-44.
- Graur, D. and W.-H. Li (2000). Fundamentals of Molecular Evolution. Sunderland, Massachusetts, Sinauer Associates, Inc.
- Gray, D. M. (1997). "Derivation of nearest-neighbor properties from data on nucleic acid oligomers. II. Thermodynamic parameters of DNA.RNA hybrids and DNA duplexes." Biopolymers **42**(7): 795-810.
- Griffiths-Jones, S., A. Bateman, et al. (2003). "Rfam: an RNA family database." Nucleic Acids

- Res **31**(1): 439-41.
- Griffiths-Jones, S., R. J. Grocock, et al. (2006). "miRBase: microRNA sequences, targets and gene nomenclature." Nucleic Acids Res **34**(Database issue): D140-4.
- Griffiths-Jones, S., S. Moxon, et al. (2005). "Rfam: annotating non-coding RNAs in complete genomes." Nucleic Acids Res **33 Database Issue**: D121-4.
- Grosshans, H. and F. J. Slack (2002). "Micro-RNAs: small is plentiful." J Cell Biol **156**(1): 17-21.
- Gu, X., Y. Wang, et al. (2002). "Age distribution of human gene families shows significant roles of both large- and small-scale duplications in vertebrate evolution." Nat Genet **31**(2): 205-9.
- GuhaThakurta, D. and G. D. Stormo (2001). "Identifying target sites for cooperatively binding factors." Bioinformatics **17**(7): 608-21.
- Gunnery, S., Y. Ma, et al. (1999). "Termination sequence requirements vary among genes transcribed by RNA polymerase III." J Mol Biol **286**(3): 745-57.
- Gutell, R. R., N. Larsen, et al. (1994). "Lessons from an evolving rRNA: 16S and 23S rRNA structures from a comparative perspective." Microbiol Rev **58**(1): 10-26.
- Gutell, R. R., A. Power, et al. (1992). "Identifying constraints on the higher-order structure of RNA: continued development and application of comparative sequence analysis methods." Nucleic Acids Res **20**(21): 5785-95.
- Guthrie, C. and J. Abelson (1982). Organization and Expression of tRNA Genes in *Saccharomyces cerevisiae*. The Molecular Biology of the Yeast Saccharomyces: Metabolism and Gene Expression. J. R. Broach. Woodbury, New York, Cold Spring Harbor Laboratory Press: 487-528.
- Hallenberg, C., J. Nederby Nielsen, et al. (1994). "Characterization of 5S rRNA genes from mouse." Gene **142**(2): 291-5.
- Havgaard, J. H., R. B. Lyngso, et al. (2005). "Pairwise local structural alignment of RNA sequences with sequence similarity less than 40%." Bioinformatics **21**(9): 1815-24.
- Hazkani-Covo, E., R. Sorek, et al. (2003). "Evolutionary dynamics of large numts in the human genome: rarity of independent insertions and abundance of post-insertion duplications." J Mol Evol **56**(2): 169-74.
- Helm, M. (2006). "Post-transcriptional nucleotide modification and alternative folding of RNA." Nucleic Acids Res **34**(2): 721-33.
- Hermann, T. and E. Westhof (1999). "Non-Watson-Crick base pairs in RNA-protein recognition." Chem Biol **6**(12): R335-43.
- Hillier, L. W., W. Miller, et al. (2004). "Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution." Nature **432**(7018): 695-716.
- Hochsmann, M., B. Voss, et al. (2004). "Pure multiple RNA secondary structure alignments: a

- progressive profile approach." *IEEE/ACM Trans Comput Biol Bioinform* **1**(1): 53-62.
- Hofacker, I. L. (2003). "Vienna RNA secondary structure server." *Nucleic Acids Res* **31**(13): 3429-31.
- Hofacker, I. L. (2006). "RNAplot." <http://www.tbi.univie.ac.at/~ivo/RNA/RNAplot.html>.
- Hofacker, I. L., S. H. Bernhart, et al. (2004). "Alignment of RNA base pairing probability matrices." *Bioinformatics* **20**(14): 2222-7.
- Hofacker, I. L., M. Fekete, et al. (2002). "Secondary structure prediction for aligned RNA sequences." *J Mol Biol* **319**(5): 1059-66.
- Hofacker, I. L., W. Fontana, et al. (1994-2006). "RNAfold." <http://www.tbi.univie.ac.at/~ivo/RNA/RNAfold.html>.
- Hoffmann, A. A., C. M. Sgro, et al. (2004). "Chromosomal inversion polymorphisms and adaptation." *Trends Ecol Evol* **19**(9): 482-8.
- Hsieh, Y. J., Z. Wang, et al. (1999). "Cloning and characterization of two evolutionarily conserved subunits (TFIIIC102 and TFIIIC63) of human TFIIIC and their involvement in functional interactions with TFIIIB and RNA polymerase III." *Mol Cell Biol* **19**(7): 4944-52.
- Hu, J., B. Li, et al. (2005). "Limitations and potentials of current motif discovery algorithms." *Nucleic Acids Res* **33**(15): 4899-913.
- Hu, Y. J. (2002). "Prediction of consensus structural motifs in a family of coregulated RNA sequences." *Nucleic Acids Res* **30**(17): 3886-93.
- Huang, X. and A. Madan (1999). "CAP3: A DNA sequence assembly program." *Genome Res* **9**(9): 868-77.
- Ingham, C. J., J. Dennis, et al. (1999). "Autogenous regulation of transcription termination factor Rho and the requirement for Nus factors in *Bacillus subtilis*." *Mol Microbiol* **31**(2): 651-63.
- International Human Genome Sequencing Consortium (2001). "Initial sequencing and analysis of the human genome." *Nature* **409**(6822): 860-921.
- James, B. D., G. J. Olsen, et al. (1989). "Phylogenetic comparative analysis of RNA secondary structure." *Methods Enzymol* **180**: 227-39.
- Ji, Y., X. Xu, et al. (2004). "A graph theoretical approach for predicting common RNA secondary structure motifs including pseudoknots in unaligned sequences." *Bioinformatics* **20**(10): 1591-602.
- Joyce, G. F. (2002). "The antiquity of RNA-based evolution." *Nature* **418**(6894): 214-21.
- Kampa, D., J. Cheng, et al. (2004). "Novel RNAs identified from an in-depth analysis of the transcriptome of human chromosomes 21 and 22." *Genome Res* **14**(3): 331-42.
- Kanz, C., P. Aldebert, et al. (2005). "The EMBL Nucleotide Sequence Database." *Nucleic Acids Res* **33**(Database issue): D29-33.
- Katayama, S., Y. Tomaru, et al. (2005). "Antisense transcription in the mammalian

- transcriptome." *Science* **309**(5740): 1564-6.
- Ke, A. and J. A. Doudna (2004). "Crystallization of RNA and RNA-protein complexes." *Methods* **34**(3): 408-14.
- Khoury, G. and P. Gruss (1983). "Enhancer elements." *Cell* **33**(2): 313-4.
- Kiss, T. and W. Filipowicz (1995). "Exonucleolytic processing of small nucleolar RNAs from pre-mRNA introns." *Genes Dev* **9**(11): 1411-24.
- Klein, R. J. and S. R. Eddy (2003). "RSEARCH: finding homologs of single structured RNA sequences." *BMC Bioinformatics* **4**: 44.
- Klosterman, P. S., D. K. Hendrix, et al. (2004). "Three-dimensional motifs from the SCOR, structural classification of RNA database: extruded strands, base triples, tetraloops and U-turns." *Nucleic Acids Res* **32**(8): 2342-52.
- Knudsen, B. and J. Hein (1999). "RNA secondary structure prediction using stochastic context-free grammars and evolutionary history." *Bioinformatics* **15**(6): 446-54.
- Knudsen, B. and J. Hein (2003). "Pfold: RNA secondary structure prediction using stochastic context-free grammars." *Nucleic Acids Res* **31**(13): 3423-8.
- Konings, D. A. and R. R. Gutell (1995). "A comparison of thermodynamic foldings with comparatively derived structures of 16S and 16S-like rRNAs." *Rna* **1**(6): 559-74.
- Kozak, M. (2005). "Regulation of translation via mRNA structure in prokaryotes and eukaryotes." *Gene* **361**: 13-37.
- Kramerov, D. A. and N. S. Vassetzky (2005). "Short retroposons in eukaryotic genomes." *Int Rev Cytol* **247**: 165-221.
- Kruger, K., P. J. Grabowski, et al. (1982). "Self-splicing RNA: autoexcision and autocyclization of the ribosomal RNA intervening sequence of Tetrahymena." *Cell* **31**(1): 147-57.
- Lasser-Weiss, M., N. Bawnik, et al. (1981). "Isolation and characterization of cloned rat DNA fragment carrying tRNA genes." *Nucleic Acids Res* **9**(22): 5965-78.
- Lawn, R. M., K. Schwartz, et al. (1997). "Convergent evolution of apolipoprotein(a) in primates and hedgehog." *Proc Natl Acad Sci U S A* **94**(22): 11992-7.
- Lawrence, C. E., S. F. Altschul, et al. (1993). "Detecting subtle sequence signals: a Gibbs sampling strategy for multiple alignment." *Science* **262**(5131): 208-14.
- Lawrence, C. E. and A. A. Reilly (1990). "An expectation maximization (EM) algorithm for the identification and characterization of common sites in unaligned biopolymer sequences." *Proteins* **7**(1): 41-51.
- Lee, Y., C. Ahn, et al. (2003). "The nuclear RNase III Drosha initiates microRNA processing." *Nature* **425**(6956): 415-9.
- Lee, Y., M. Kim, et al. (2004). "MicroRNA genes are transcribed by RNA polymerase II." *Embo J* **23**(20): 4051-60.
- Leontis, N. B. and E. Westhof (2003). "Analysis of RNA motifs." *Curr Opin Struct Biol* **13**(3):

300-8.

- Lesnik, E. A., R. Sampath, et al. (2001). "Prediction of rho-independent transcriptional terminators in *Escherichia coli*." *Nucleic Acids Res* **29**(17): 3583-94.
- Li, Q., K. R. Peterson, et al. (2002). "Locus control regions." *Blood* **100**(9): 3077-86.
- Lindgreen, S., P. P. Gardner, et al. (2006). "Measuring covariation in RNA alignments: physical realism improves information measures." *Bioinformatics* **22**(24): 2988-95.
- Little, R. D. and D. C. Braaten (1989). "Genomic organization of human 5 S rDNA and sequence of one tandem repeat." *Genomics* **4**(3): 376-83.
- Liu, X., D. L. Brutlag, et al. (2001). "BioProspector: discovering conserved DNA motifs in upstream regulatory regions of co-expressed genes." *Pac Symp Biocomput*: 127-38.
- Lopez-Lastra, M., A. Rivas, et al. (2005). "Protein synthesis in eukaryotes: the growing biological relevance of cap-independent translation initiation." *Biol Res* **38**(2-3): 121-46.
- Lowe, T. M. and S. R. Eddy (1997). "tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence." *Nucleic Acids Res* **25**(5): 955-64.
- Lukavsky, P. J. and J. D. Puglisi (2005). "Structure determination of large biological RNAs." *Methods Enzymol* **394**: 399-416.
- Lynch, M. and J. S. Conery (2000). "The evolutionary fate and consequences of duplicate genes." *Science* **290**(5494): 1151-5.
- Lyngso, R. B. and C. N. Pedersen (2000). "RNA pseudoknot prediction in energy-based models." *J Comput Biol* **7**(3-4): 409-27.
- MacIsaac, K. D. and E. Fraenkel (2006). "Practical strategies for discovering regulatory DNA sequence motifs." *PLoS Comput Biol* **2**(4): e36.
- Maestre, J., T. Tchenio, et al. (1995). "mRNA retroposition in human cells: processed pseudogene formation." *Embo J* **14**(24): 6333-8.
- Mandal, M., B. Boese, et al. (2003). "Riboswitches control fundamental biochemical pathways in *Bacillus subtilis* and other bacteria." *Cell* **113**(5): 577-86.
- Marsan, L. and M. F. Sagot (2000). "Algorithms for extracting structured motifs using a suffix tree with an application to promoter and regulatory site consensus identification." *J Comput Biol* **7**(3-4): 345-62.
- Matouk, I. J., N. DeGroot, et al. (2007). "The H19 non-coding RNA is essential for human tumor growth." *PLoS ONE* **2**(9): e845.
- Matsui, H., K. Sato, et al. (2004). "Pair stochastic tree adjoining grammars for aligning and predicting pseudoknot RNA structures." *Proc IEEE Comput Syst Bioinform Conf*: 290-9.
- Mattaj, I. W. (1993). "RNA recognition: a family matter?" *Cell* **73**(5): 837-40.
- McCaskill, J. S. (1990). "The equilibrium partition function and base pair binding probabilities for RNA secondary structure." *Biopolymers* **29**(6-7): 1105-19.

- McCullagh, P. and J. A. Nelder (1983). Generalized linear models, Chapman and Hall, London.
- Morlando, M., P. Greco, et al. (2002). "Functional analysis of yeast snoRNA and snRNA 3'-end formation mediated by uncoupling of cleavage and polyadenylation." Mol Cell Biol **22**(5): 1379-89.
- Mourier, T., A. J. Hansen, et al. (2001). "The Human Genome Project reveals a continuous transfer of large mitochondrial fragments to the nucleus." Mol Biol Evol **18**(9): 1833-7.
- Mouse Genome Sequencing Consortium (2002). "Initial sequencing and comparative analysis of the mouse genome." Nature **420**(6915): 520-62.
- Mungall, A. J., S. A. Palmer, et al. (2003). "The DNA sequence and analysis of human chromosome 6." Nature **425**(6960): 805-11.
- Nahvi, A., J. E. Barrick, et al. (2004). "Coenzyme B12 riboswitches are widespread genetic control elements in prokaryotes." Nucleic Acids Res **32**(1): 143-50.
- Nakanishi, K. and O. Nureki (2005). "Recent progress of structural biology of tRNA processing and modification." Mol Cells **19**(2): 157-66.
- Nam, J. W., K. R. Shin, et al. (2005). "Human microRNA prediction through a probabilistic co-learning model of sequence and structure." Nucleic Acids Res **33**(11): 3570-81.
- Neidle, S. (2002). Nucleic acid structure and recognition, Oxford University Press, New York.
- Nissen, P., J. Hansen, et al. (2000). "The structural basis of ribosome activity in peptide bond synthesis." Science **289**(5481): 920-30.
- Nomenclature Committee of the International Union of Biochemistry, N.-I. (1986). "Nomenclature for incompletely specified bases in nucleic acid sequences. Recommendations 1984." Proc Natl Acad Sci U S A **83**(1): 4-8.
- Notredame, C., D. G. Higgins, et al. (2000). "T-Coffee: A novel method for fast and accurate multiple sequence alignment." J Mol Biol **302**(1): 205-17.
- Nussinov, R. and A. B. Jacobson (1980). "Fast algorithm for predicting the secondary structure of single-stranded RNA." Proc Natl Acad Sci U S A **77**(11): 6903-13.
- Ohler, U., S. Yekta, et al. (2004). "Patterns of flanking sequence conservation and a characteristic upstream motif for microRNA gene identification." Rna **10**(9): 1309-22.
- Ohshima, K., M. Hattori, et al. (2003). "Whole-genome screening indicates a possible burst of formation of processed pseudogenes and Alu repeats by particular L1 subfamilies in ancestral primates." Genome Biol **4**(11): R74.
- Okazaki, Y., M. Furuno, et al. (2002). "Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs." Nature **420**(6915): 563-73.
- Olson, T., M. J. Fournier, et al. (1976). "Detection of a major conformational change in transfer ribonucleic acid by laser light scattering." J Mol Biol **102**(2): 193-203.
- Onoa, B. and I. Tinoco, Jr. (2004). "RNA folding and unfolding." Curr Opin Struct Biol **14**(3): 374-9.

- Osada, R., E. Zaslavsky, et al. (2004). "Comparative analysis of methods for representing and searching for transcription factor binding sites." *Bioinformatics* **20**(18): 3516-25.
- Ota, T., Y. Suzuki, et al. (2004). "Complete sequencing and characterization of 21,243 full-length human cDNAs." *Nat Genet* **36**(1): 40-5.
- Pasquinelli, A. E., B. J. Reinhart, et al. (2000). "Conservation of the sequence and temporal expression of let-7 heterochronic regulatory RNA." *Nature* **408**(6808): 86-9.
- Passananti, C., B. Davies, et al. (1987). "Structure of an inverted duplication formed as a first step in a gene amplification event: implications for a model of gene amplification." *Embo J* **6**(6): 1697-703.
- Paule, M. R. and R. J. White (2000). "Survey and summary: transcription by RNA polymerases I and III." *Nucleic Acids Res* **28**(6): 1283-98.
- Pavesi, A., F. Conterio, et al. (1994). "Identification of new eukaryotic tRNA genes in genomic DNA databases by a multistep weight matrix analysis of transcriptional control regions." *Nucleic Acids Res* **22**(7): 1247-56.
- Pavesi, G., G. Mauri, et al. (2001). "An algorithm for finding signals of unknown length in DNA sequences." *Bioinformatics* **17 Suppl 1**: S207-14.
- Pavesi, G., G. Mauri, et al. (2004). "In silico representation and discovery of transcription factor binding sites." *Brief Bioinform* **5**(3): 217-36.
- Pedersen, J. S., G. Bejerano, et al. (2006). "Identification and classification of conserved RNA secondary structures in the human genome." *PLoS Comput Biol* **2**(4): e33.
- Ramadass, A. S. (2004). Computational detection of gene regulatory signals in human genome sequence. *The Wellcome Trust Sanger Institute*. Hinxton, Cambridge, Cambridge.
- Rangan, P. and S. A. Woodson (2003). "Structural requirement for Mg<sup>2+</sup> binding in the group I intron core." *J Mol Biol* **329**(2): 229-38.
- Ravasi, T., H. Suzuki, et al. (2006). "Experimental validation of the regulated expression of large numbers of non-coding RNAs from the mouse genome." *Genome Res* **16**(1): 11-9.
- Reams, A. B. and E. L. Neidle (2004). "Selection for gene clustering by tandem duplication." *Annu Rev Microbiol* **58**: 119-42.
- Reddy, R., D. Henning, et al. (1987). "The capped U6 small nuclear RNA is transcribed by RNA polymerase III." *J Biol Chem* **262**(1): 75-81.
- Ricchetti, M., F. Tekaiia, et al. (2004). "Continued colonization of the human genome by mitochondrial DNA." *PLoS Biol* **2**(9): E273.
- Riek, R., K. Pervushin, et al. (2000). "TROSY and CRINEPT: NMR with large molecular and supramolecular structures in solution." *Trends Biochem Sci* **25**(10): 462-8.
- Rietveld, K., R. Van Poelgeest, et al. (1982). "The tRNA-like structure at the 3' terminus of turnip yellow mosaic virus RNA. Differences and similarities with canonical tRNA." *Nucleic Acids Res* **10**(6): 1929-46.

- Ringner, M. and M. Krogh (2005). "Folding Free Energies of 5'-UTRs Impact Post-Transcriptional Regulation on a Genomic Scale in Yeast." *PLoS Comput Biol* **1**(7): e72.
- Rivas, E. and S. R. Eddy (1999). "A dynamic programming algorithm for RNA structure prediction including pseudoknots." *J Mol Biol* **285**(5): 2053-68.
- Rivas, E. and S. R. Eddy (2000). "The language of RNA: a formal grammar that includes pseudoknots." *Bioinformatics* **16**(4): 334-40.
- Rivas, E. and S. R. Eddy (2000). "Secondary structure alone is generally not statistically significant for the detection of noncoding RNAs." *Bioinformatics* **16**(7): 583-605.
- Rivas, E. and S. R. Eddy (2001). "Noncoding RNA gene detection using comparative sequence analysis." *BMC Bioinformatics* **2**(1): 8.
- Rivas, E., R. J. Klein, et al. (2001). "Computational identification of noncoding RNAs in *E. coli* by comparative genomics." *Curr Biol* **11**(17): 1369-73.
- Rogic, S., A. K. Mackworth, et al. (2001). "Evaluation of gene-finding programs on mammalian sequences." *Genome Res* **11**(5): 817-32.
- Romero, D. and R. Palacios (1997). "Gene amplification and genomic plasticity in prokaryotes." *Annu Rev Genet* **31**: 91-111.
- Rould, M. A., J. J. Perona, et al. (1991). "Structural basis of anticodon loop recognition by glutamyl-tRNA synthetase." *Nature* **352**(6332): 213-8.
- Ruan, J., G. D. Stormo, et al. (2004). "An iterated loop matching approach to the prediction of RNA secondary structures with pseudoknots." *Bioinformatics* **20**(1): 58-66.
- Sakakibara, Y., M. Brown, et al. (1994). "Stochastic context-free grammars for tRNA modeling." *Nucleic Acids Res* **22**(23): 5112-20.
- Sandelin, A., P. Carninci, et al. (2007). "Mammalian RNA polymerase II core promoters: insights from genome-wide studies." *Nat Rev Genet* **8**(6): 424-36.
- Sankoff, D. (1985). "Simultaneous solution of the RNA folding, alignment and protosequence problems." *SIAM J. Appl. Math.* **45**(5): 810-825.
- Sankoff, D. and N. El-Mabrouk (2000). Genome Rearrangement. *Current topics in computational biology*. M. Zhang. Cambridge, MIT Press: 135-155.
- Schmeing, T. M., A. C. Seila, et al. (2002). "A pre-translocational intermediate in protein synthesis observed in crystals of enzymatically active 50S subunits." *Nat Struct Biol* **9**(3): 225-30.
- Schneider, T. D. and R. M. Stephens (1990). "Sequence logos: a new way to display consensus sequences." *Nucleic Acids Res* **18**(20): 6097-100.
- Schwartz, S., W. J. Kent, et al. (2003). "Human-mouse alignments with BLASTZ." *Genome Res* **13**(1): 103-7.
- Seitz, H., H. Royo, et al. (2004). "A large imprinted microRNA gene cluster at the mouse *Dlk1-Gtl2* domain." *Genome Res* **14**(9): 1741-8.

- Shanab, G. M. and E. S. Maxwell (1991). "Proposed secondary structure of eukaryotic U14 snRNA." *Nucleic Acids Res* **19**(18): 4891-4.
- Siebert, S. and R. Backofen (2005). "MARNA: multiple alignment and consensus structure prediction of RNAs based on sequence structure comparisons." *Bioinformatics* **21**(16): 3352-9.
- Siepel, A., G. Bejerano, et al. (2005). "Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes." *Genome Res* **15**(8): 1034-50.
- Sinha, S. and M. Tompa (2000). "A statistical method for finding transcription factor binding sites." *Proc Int Conf Intell Syst Mol Biol* **8**: 344-54.
- Sinha, S. and M. Tompa (2002). "Discovery of novel transcription factor binding sites by statistical overrepresentation." *Nucleic Acids Res* **30**(24): 5549-60.
- Slack, F. J., M. Basson, et al. (2000). "The lin-41 RBCC gene acts in the C. elegans heterochronic pathway between the let-7 regulatory RNA and the LIN-29 transcription factor." *Mol Cell* **5**(4): 659-69.
- Smit, A. F. (1999). "Interspersed repeats and other mementos of transposable elements in mammalian genomes." *Curr Opin Genet Dev* **9**(6): 657-63.
- Smit, A. F. and A. D. Riggs (1995). "MIRs are classic, tRNA-derived SINEs that amplified before the mammalian radiation." *Nucleic Acids Res* **23**(1): 98-102.
- Smit, A. F. A. and P. Green (unpublished). "RepeatMasker."  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>.
- Sprinzi, M., C. Horn, et al. (1998). "Compilation of tRNA sequences and sequences of tRNA genes." *Nucleic Acids Res* **26**(1): 148-53.
- Sprinzi, M. and K. S. Vassilenko (2005). "Compilation of tRNA sequences and sequences of tRNA genes." *Nucleic Acids Res* **33**(Database issue): D139-40.
- Staple, D. W. and S. E. Butcher (2005). "Pseudoknots: RNA structures with diverse functions." *PLoS Biol* **3**(6): e213.
- Steege, D. A. (2000). "Emerging features of mRNA decay in bacteria." *Rna* **6**(8): 1079-90.
- Stormo, G. D. (2000). "DNA binding sites: representation and discovery." *Bioinformatics* **16**(1): 16-23.
- Storz, G. (2002). "An expanding universe of noncoding RNAs." *Science* **296**(5571): 1260-3.
- Suzuki, H., K. Moriwaki, et al. (1994). "Sequences and evolutionary analysis of mouse 5S rDNAs." *Mol Biol Evol* **11**(4): 704-10.
- Tamura, M., D. K. Hendrix, et al. (2004). "SCOR: Structural Classification of RNA, version 2.0." *Nucleic Acids Res* **32**(Database issue): D182-4.
- Taneda, A. (2005). "Cofolga: a genetic algorithm for finding the common folding of two RNAs." *Comput Biol Chem* **29**(2): 111-9.
- The ENCODE Project Consortium (2007). "Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project." *Nature* **447**(7146):

- 799-816.
- Thompson, W., E. C. Rouchka, et al. (2003). "Gibbs Recursive Sampler: finding transcription factor binding sites." Nucleic Acids Res **31**(13): 3580-5.
- TIGR (2002-2003). "TIGR Gene Indices Clustering Tools (TGICL)."
- Tipping, M. E. (1999). "The relevance vector machine." Advances in Neural Information Processing Systems 12 [NIPS Conference, Denver, Colorado, USA, November 29 - December 4, 1999]: 652-658.
- Tompa, M. (1999). "An exact method for finding short motifs in sequences, with application to the ribosome binding site problem." Proc Int Conf Intell Syst Mol Biol: 262-71.
- Tompa, M., N. Li, et al. (2005). "Assessing computational tools for the discovery of transcription factor binding sites." Nat Biotechnol **23**(1): 137-44.
- Torarinsson, E., M. Sawera, et al. (2006). "Thousands of corresponding human and mouse genomic regions unalignable in primary sequence contain common RNA structure." Genome Res **16**(7): 885-9.
- Torrents, D., M. Suyama, et al. (2003). "A genome-wide survey of human pseudogenes." Genome Res **13**(12): 2559-67.
- Tourmen, Y., O. Baris, et al. (2002). "Structure and chromosomal distribution of human mitochondrial pseudogenes." Genomics **80**(1): 71-7.
- Tsuzuki, T., H. Nomiya, et al. (1983). "Presence of mitochondrial-DNA-like sequences in the human nuclear DNA." Gene **25**(2-3): 223-9.
- Uemura, Y., A. Hasegawa, et al. (1999). "Tree adjoining grammars for RNA structure prediction." Theoretical Computer Science **210**(2): 277-303.
- Uptain, S. M. and M. J. Chamberlin (1997). "Escherichia coli RNA polymerase terminates transcription efficiently at rho-independent terminators on single-stranded DNA templates." Proc Natl Acad Sci U S A **94**(25): 13548-53.
- van Batenburg, F. H., A. P. Gulyaev, et al. (2001). "PseudoBase: structural information on RNA pseudoknots." Nucleic Acids Res **29**(1): 194-5.
- Van de Peer, Y. (2004). "Computational approaches to unveiling ancient genome duplications." Nat Rev Genet **5**(10): 752-63.
- Varani, G. and A. Pardi (1994). Structure of RNA. RNA-Protein Interactions. M. I.W. Oxford University, Oxford, IRL PRESS: 1-24.
- Vignali, M., A. H. Hassan, et al. (2000). "ATP-dependent chromatin-remodeling complexes." Mol Cell Biol **20**(6): 1899-910.
- Walter, A. E., D. H. Turner, et al. (1994). "Coaxial stacking of helices enhances binding of oligoribonucleotides and improves predictions of RNA folding." Proc Natl Acad Sci U S A **91**(20): 9218-22.
- Washietl, S., I. L. Hofacker, et al. (2005). "Mapping of conserved RNA secondary structures predicts thousands of functional noncoding RNAs in the human genome." Nat

- Biotechnol **23**(11): 1383-90.
- Washietl, S., I. L. Hofacker, et al. (2005). "Fast and reliable prediction of noncoding RNAs." Proc Natl Acad Sci U S A **102**(7): 2454-9.
- Washietl, S., J. S. Pedersen, et al. (2007). "Structured RNAs in the ENCODE selected regions of the human genome." Genome Res **17**(6): 852-64.
- Wasserman, W. W. and J. W. Fickett (1998). "Identification of regulatory regions which confer muscle-specific gene expression." J Mol Biol **278**(1): 167-81.
- Wasserman, W. W. and A. Sandelin (2004). "Applied bioinformatics for the identification of regulatory elements." Nat Rev Genet **5**(4): 276-87.
- Weinmann, R., H. J. Raskas, et al. (1974). "Role of DNA-dependent RNA polymerases II and III in transcription of the adenovirus genome late in productive infection." Proc Natl Acad Sci U S A **71**(9): 3426-39.
- Weischenfeldt, J., J. Lykke-Andersen, et al. (2005). "Messenger RNA surveillance: neutralizing natural nonsense." Curr Biol **15**(14): R559-62.
- Westhof, E. and F. Michel (1994). Prediction and experimental investigation of RNA secondary and tertiary foldings. RNA-Protein Interactions. M. I.W. Oxford University, Oxford, IRL PRESS: 26-51.
- Will, C. L. and R. Luhrmann (2001). "Spliceosomal UsnRNP biogenesis, structure and function." Curr Opin Cell Biol **13**(3): 290-301.
- Wilson, J. H., J. S. Kim, et al. (1972). "Bacteriophage T4 transfer RNA. 3. Clustering of the genes for the T4 transfer RNA's." J Mol Biol **71**(3): 547-56.
- Woischnik, M. and C. T. Moraes (2002). "Pattern of organization of human mitochondrial pseudogenes in the nuclear genome." Genome Res **12**(6): 885-93.
- Wu, G. J., J. F. Railey, et al. (1987). "Defining the functional domains in the control region of the adenovirus type 2 specific VARNA1 gene." J Mol Biol **194**(3): 423-42.
- Xue, C., F. Li, et al. (2005). "Classification of real and pseudo microRNA precursors using local structure-sequence features and support vector machine." BMC Bioinformatics **6**: 310.
- Zhang, Z., P. M. Harrison, et al. (2003). "Millions of years of evolution preserved: a comprehensive catalog of the processed pseudogenes in the human genome." Genome Res **13**(12): 2541-58.
- Zhou, Q. and W. H. Wong (2004). "CisModule: de novo discovery of cis-regulatory modules by hierarchical mixture modeling." Proc Natl Acad Sci U S A **101**(33): 12114-9.
- Zuker, M. (1989). "On finding all suboptimal foldings of an RNA molecule." Science **244**(4900): 48-52.
- Zuker, M. and P. Stiegler (1981). "Optimal computer folding of large RNA sequences using thermodynamics and auxiliary information." Nucleic Acids Res **9**(1): 133-48.