

References

- (1998). "The C. elegans Sequencing Consortium. (1998) Genome sequence of the nematode C. elegans: a platform for investigating biology." Science **282**(5396): 2012-8.
- Alberts, B., D. Bray, et al. (2002). Molecular Biology of the Cell, Garland.
- Alt, F. W., V. Enea, et al. (1980). "Activity of multiple light chain genes in murine myeloma cells producing a single, functional light chain." Cell **21**(1): 1-12.
- Awasthi, S. and J. C. Alwine (2003). "Association of polyadenylation cleavage factor I with U1 snRNP." Rna **9**(11): 1400-9.
- 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.
- Bateman, A., L. Coin, et al. (2004). "The Pfam protein families database." Nucleic Acids Res **32 Database issue**: D138-41.
- Beaudoing, E., S. Freier, et al. (2000). "Patterns of variant polyadenylation signal usage in human genes." Genome Res **10**(7): 1001-10.
- Berget, S. M. (1995). "Exon recognition in vertebrate splicing." J Biol Chem **270**(6): 2411-4.
- Bernstein, D. S., N. Buter, et al. (2002). "Analyzing mRNA-protein complexes using a yeast three-hybrid system." Methods **26**(2): 123-41.
- Bernstein, D. S., B. Hook, et al. (2005). "Binding specificity and mRNA targets of a C. elegans PUF protein, FBF-1." RNA **11**(4): 447-58.
- Bienroth, S., W. Keller, et al. (1993). "Assembly of a processive messenger RNA polyadenylation complex." Embo J **12**(2): 585-94.
- Blumenthal, T. (1995). "Trans-splicing and polycistronic transcription in Caenorhabditis elegans." Trends Genet **11**(4): 132-6.
- Blumenthal, T. and K. Steward (1997). RNA processing and Gene Structure. C. elegans II. D. Riddle, T. Blumenthal, B. Meyer and J. R. Preiss. Cold Spring Harbour, Cold Spring Harbour Laboratory Press: 117-145.
- Bousquet-Antonelli, C., C. Presutti, et al. (2000). "Identification of a regulated pathway for nuclear pre-mRNA turnover." Cell **102**(6): 765-75.
- Brukner, I., R. Sanchez, et al. (1995). "Sequence-dependent bending propensity of DNA as revealed by DNase I: parameters for trinucleotides." Embo J **14**(8): 1812-8.
- Burge, C. and S. Karlin (1997). "Prediction of complete gene structures in human genomic DNA." J Mol Biol **268**(1): 78-94.
- Cardinali, B., M. Di Cristina, et al. (1993). "Interaction of proteins with the mRNA for ribosomal protein L1 in Xenopus: structural characterization of in vivo complexes and identification of proteins that bind in vitro to its 5'UTR." Nucleic Acids Res **21**(10): 2301-8.
- Castelo-Branco, P., A. Furger, et al. (2004). "Polypyrimidine tract binding protein modulates efficiency of polyadenylation." Mol Cell Biol **24**(10): 4174-83.
- Chen, C. Y. and A. B. Shyu (2003). "Rapid deadenylation triggered by a nonsense codon precedes decay of the RNA body in a mammalian cytoplasmic nonsense-mediated decay pathway." Mol Cell Biol **23**(14): 4805-13.

- Chen, F., C. C. MacDonald, et al. (1995). "Cleavage site determinants in the mammalian polyadenylation signal." *Nucleic Acids Res* **23**(14): 2614-20.
- Chen, F. and J. Wilusz (1998). "Auxiliary downstream elements are required for efficient polyadenylation of mammalian pre-mRNAs." *Nucleic Acids Res* **26**(12): 2891-8.
- Chen, N., T. W. Harris, et al. (2005). "WormBase: a comprehensive data resource for *Caenorhabditis* biology and genomics." *Nucleic Acids Res* **33 Database Issue**: D383-9.
- Clamp, M., J. Cuff, et al. (2004). "The Jalview Java alignment editor." *Bioinformatics* **20**(3): 426-7.
- Colgan, D. F. and J. L. Manley (1997). "Mechanism and regulation of mRNA polyadenylation." *Genes Dev* **11**(21): 2755-66.
- Cooke, C. and J. C. Alwine (1996). "The cap and the 3' splice site similarly affect polyadenylation efficiency." *Mol Cell Biol* **16**(6): 2579-84.
- Dantoni, J. C., K. G. Murthy, et al. (1997). "Transcription factor TFIID recruits factor CPSF for formation of 3' end of mRNA." *Nature* **389**(6649): 399-402.
- Darnell, J. E., R. Wall, et al. (1971). "An adenylic acid-rich sequence in messenger RNA of HeLa cells and its possible relationship to reiterated sites in DNA." *Proc Natl Acad Sci U S A* **68**(6): 1321-5.
- Dermitzakis, E. T., C. M. Bergman, et al. (2003). "Tracing the evolutionary history of *Drosophila* regulatory regions with models that identify transcription factor binding sites." *Mol Biol Evol* **20**(5): 703-14.
- Dichtl, B. and W. Keller (2001). "Recognition of polyadenylation sites in yeast pre-mRNAs by cleavage and polyadenylation factor." *Embo J* **20**(12): 3197-209.
- Dominski, Z. and W. F. Marzluff (1999). "Formation of the 3' end of histone mRNA." *Gene* **239**(1): 1-14.
- 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.
- Dsouza, M., N. Larsen, et al. (1997). "Searching for patterns in genomic data." *Trends Genet* **13**(12): 497-8.
- Durbin, R., S. R. Eddy, et al. (1998). *Biological Sequence Analysis*. Cambridge, Cambridge University Press.
- Eddy, S. R. (2001). "Non-coding RNA genes and the modern RNA world." *Nat Rev Genet* **2**(12): 919-29.
- Eddy, S. R. (2002). "A memory-efficient dynamic programming algorithm for optimal alignment of a sequence to an RNA secondary structure." *BMC Bioinformatics* **3**(1): 18.
- Edmonds, M., M. H. Vaughan, Jr., et al. (1971). "Polyadenylic acid sequences in the heterogeneous nuclear RNA and rapidly-labeled polyribosomal RNA of HeLa cells: possible evidence for a precursor relationship." *Proc Natl Acad Sci U S A* **68**(6): 1336-40.
- Edwards-Gilbert, G., K. L. Veraldi, et al. (1997). "Alternative poly(A) site selection in complex transcription units: means to an end?" *Nucleic Acids Res* **25**(13): 2547-61.
- Evans, D., I. Perez, et al. (2001). "A complex containing CstF-64 and the SL2 snRNP connects mRNA 3' end formation and trans-splicing in *C. elegans* operons." *Genes Dev* **15**(19): 2562-71.

- Ford, J. P. and M. T. Hsu (1978). "Transcription pattern of in vivo-labeled late simian virus 40 RNA: equimolar transcription beyond the mRNA 3' terminus." J Virol **28**(3): 795-801.
- Ford, L. P., P. S. Bagga, et al. (1997). "The poly(A) tail inhibits the assembly of a 3'-to-5' exonuclease in an in vitro RNA stability system." Mol Cell Biol **17**(1): 398-406.
- Fromont-Racine, M., B. Senger, et al. (2003). "Ribosome assembly in eukaryotes." Gene **313**: 17-42.
- Gardiner-Garden, M. and M. Frommer (1987). "CpG islands in vertebrate genomes." J Mol Biol **196**(2): 261-82.
- 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.
- Gavis, E. R., D. Curtis, et al. (1996). "Identification of cis-acting sequences that control nanos RNA localization." Dev Biol **176**(1): 36-50.
- Gill, G. (2001). "Regulation of the initiation of eukaryotic transcription." Essays Biochem **37**: 33-43.
- Graber, J. H., C. R. Cantor, et al. (1999). "In silico detection of control signals: mRNA 3'-end-processing sequences in diverse species." Proc Natl Acad Sci U S A **96**(24): 14055-60.
- Graber, J. H., G. D. McAllister, et al. (2002). "Probabilistic prediction of *Saccharomyces cerevisiae* mRNA 3'-processing sites." Nucleic Acids Res **30**(8): 1851-8.
- Gray, N. K. (1998). "Translational control by repressor proteins binding to the 5'UTR of mRNAs." Methods Mol Biol **77**: 379-97.
- Griffiths-Jones, S., S. Moxon, et al. (2005). "Rfam: annotating non-coding RNAs in complete genomes." Nucleic Acids Res **33 Database Issue**: D121-4.
- Gross, S. and C. L. Moore (2001). "Rna15 interaction with the A-rich yeast polyadenylation signal is an essential step in mRNA 3'-end formation." Mol Cell Biol **21**(23): 8045-55.
- Guo, Z. and F. Sherman (1996). "3'-end-forming signals of yeast mRNA." Trends Biochem Sci **21**(12): 477-81.
- Hajarnavis, A., I. Korf, et al. (2004). "A probabilistic model of 3' end formation in *Caenorhabditis elegans*." Nucleic Acids Res **32**(11): 3392-9.
- Hieronimus, H. and P. A. Silver (2004). "A systems view of mRNP biology." Genes Dev **18**(23): 2845-60.
- Higgs, D. R., S. E. Goodbourn, et al. (1983). "Alpha-thalassaemia caused by a polyadenylation signal mutation." Nature **306**(5941): 398-400.
- Hirose, Y. and J. L. Manley (1998). "RNA polymerase II is an essential mRNA polyadenylation factor." Nature **395**(6697): 93-6.
- Howe, K. L., T. Chothia, et al. (2002). "GAZE: a generic framework for the integration of gene-prediction data by dynamic programming." Genome Res **12**(9): 1418-27.
- Huang, T., S. Kuersten, et al. (2001). "Intercistronic region required for polycistronic pre-mRNA processing in *Caenorhabditis elegans*." Mol Cell Biol **21**(4): 1111-20.
- Hubbard, T., D. Andrews, et al. (2005). "Ensembl 2005." Nucleic Acids Res **33 Database Issue**: D447-53.

- Hubert, N., R. Walczak, et al. (1996). "A protein binds the selenocysteine insertion element in the 3'-UTR of mammalian selenoprotein mRNAs." Nucleic Acids Res **24**(3): 464-9.
- Jareborg, N., E. Birney, et al. (1999). "Comparative analysis of noncoding regions of 77 orthologous mouse and human gene pairs." Genome Res **9**(9): 815-24.
- Johnson, D. S., B. Davidson, et al. (2004). "Noncoding regulatory sequences of *Ciona* exhibit strong correspondence between evolutionary constraint and functional importance." Genome Res **14**(12): 2448-56.
- Kasprzyk, A., D. Keefe, et al. (2004). "EnsMart: a generic system for fast and flexible access to biological data." Genome Res **14**(1): 160-9.
- Kessler, M. M., M. F. Henry, et al. (1997). "Hrp1, a sequence-specific RNA-binding protein that shuttles between the nucleus and the cytoplasm, is required for mRNA 3'-end formation in yeast." Genes Dev **11**(19): 2545-56.
- Larizza, A., W. Makalowski, et al. (2002). "Evolutionary dynamics of mammalian mRNA untranslated regions by comparative analysis of orthologous human, artiodactyl and rodent gene pairs." Comput Chem **26**(5): 479-90.
- Lee, S. Y., J. Mendecki, et al. (1971). "A polynucleotide segment rich in adenylic acid in the rapidly-labeled polyribosomal RNA component of mouse sarcoma 180 ascites cells." Proc Natl Acad Sci U S A **68**(6): 1331-5.
- Legendre, M. and D. Gautheret (2003). "Sequence determinants in human polyadenylation site selection." BMC Genomics **4**(1): 7.
- Levy, S., D. Avni, et al. (1991). "Oligopyrimidine tract at the 5' end of mammalian ribosomal protein mRNAs is required for their translational control." Proc Natl Acad Sci U S A **88**(8): 3319-23.
- Lin, C. H. and J. G. Patton (1995). "Regulation of alternative 3' splice site selection by constitutive splicing factors." Rna **1**(3): 234-45.
- Ludwig, M. Z., C. Bergman, et al. (2000). "Evidence for stabilizing selection in a eukaryotic enhancer element." Nature **403**(6769): 564-7.
- Ludwig, M. Z., A. Palsson, et al. (2005). "Functional Evolution of a cis-Regulatory Module." PLoS Biol **3**(4): e93.
- Lund, M. and J. Kjems (2002). "Defining a 5' splice site by functional selection in the presence and absence of U1 snRNA 5' end." Rna **8**(2): 166-79.
- Lutz, C. S., K. G. Murthy, et al. (1996). "Interaction between the U1 snRNP-A protein and the 160-kD subunit of cleavage-polyadenylation specificity factor increases polyadenylation efficiency in vitro." Genes Dev **10**(3): 325-37.
- MacDonald, C. C. and J. L. Redondo (2002). "Reexamining the polyadenylation signal: were we wrong about AAUAAA?" Mol Cell Endocrinol **190**(1-2): 1-8.
- Makarov, V. (2002). "Computer programs for eukaryotic gene prediction." Brief Bioinform **3**(2): 195-9.
- Mathe, C., M. F. Sagot, et al. (2002). "Current methods of gene prediction, their strengths and weaknesses." Nucleic Acids Res **30**(19): 4103-17.
- Mattick, J. S. (2001). "Non-coding RNAs: the architects of eukaryotic complexity." EMBO Rep **2**(11): 986-91.
- Meyuhas, O. (2000). "Synthesis of the translational apparatus is regulated at the translational level." Eur J Biochem **267**(21): 6321-30.
- Mignone, F., G. Grillo, et al. (2005). "UTRdb and UTRsite: a collection of sequences and regulatory motifs of the untranslated regions of eukaryotic mRNAs." Nucleic Acids Res **33 Database Issue**: D141-6.

- Moreira, A., Y. Takagaki, et al. (1998). "The upstream sequence element of the C2 complement poly(A) signal activates mRNA 3' end formation by two distinct mechanisms." *Genes Dev* **12**(16): 2522-34.
- Morimoto, R. I. (1993). "Cells in stress: transcriptional activation of heat shock genes." *Science* **259**(5100): 1409-10.
- Neu-Yilik, G., N. H. Gehring, et al. (2004). "Nonsense-mediated mRNA decay: from vacuum cleaner to Swiss army knife." *Genome Biol* **5**(4): 218.
- Nirenberg, M. (2004). "Historical review: Deciphering the genetic code--a personal account." *Trends Biochem Sci* **29**(1): 46-54.
- Niu, D. K., K. Lin, et al. (2003). "Strand compositional asymmetries of nuclear DNA in eukaryotes." *J Mol Evol* **57**(3): 325-34.
- Niwa, M. and S. M. Berget (1991). "Mutation of the AAUAAA polyadenylation signal depresses in vitro splicing of proximal but not distal introns." *Genes Dev* **5**(11): 2086-95.
- Ohno, M., H. Sakamoto, et al. (1987). "Preferential excision of the 5' proximal intron from mRNA precursors with two introns as mediated by the cap structure." *Proc Natl Acad Sci U S A* **84**(15): 5187-91.
- Olsen, P. H. and V. Ambros (1999). "The lin-4 regulatory RNA controls developmental timing in *Caenorhabditis elegans* by blocking LIN-14 protein synthesis after the initiation of translation." *Dev Biol* **216**(2): 671-80.
- Pellizzoni, L., F. Lotti, et al. (1997). "Cellular nucleic acid binding protein binds a conserved region of the 5' UTR of *Xenopus laevis* ribosomal protein mRNAs." *J Mol Biol* **267**(2): 264-75.
- Porter, M. Y., M. Turmaine, et al. (2005). "Identification and characterization of *Caenorhabditis elegans* palmitoyl protein thioesterase 1." *J Neurosci Res* **79**(6): 836-48.
- Pritsker, M., Y. C. Liu, et al. (2004). "Whole-genome discovery of transcription factor binding sites by network-level conservation." *Genome Res* **14**(1): 99-108.
- Proudfoot, N. (1991). "Poly(A) signals." *Cell* **64**(4): 671-4.
- Proudfoot, N. J. (2001). "Genetic dangers in poly(A) signals." *EMBO Rep* **2**(10): 891-2.
- Proudfoot, N. J., A. Furger, et al. (2002). "Integrating mRNA processing with transcription." *Cell* **108**(4): 501-12.
- Qu, X., Y. Qi, et al. (2002). "Generation of multiple mRNA transcripts from the novel human apoptosis-inducing gene hap by alternative polyadenylation utilization and the translational activation function of 3' untranslated region." *Arch Biochem Biophys* **400**(2): 233-44.
- Rabiner, L. R. (1989). "A tutorial on hidden Markov models and selected applications in speech recognition." *Proceedings of the IEEE* **77**: 257-286.
- Raphael, C. (1999). "Automatic Segmentation of Acoustic Musical Signals Using Hidden Markov Models." *IEEE Transactions on Pattern Analysis and Machine Intelligence* **21**(4): 360-370.
- Rivas, E. and S. R. Eddy (2001). "Noncoding RNA gene detection using comparative sequence analysis." *BMC Bioinformatics* **2**(1): 8.
- Salamov, A. A. and V. V. Solovyev (1997). "Recognition of 3'-processing sites of human mRNA precursors." *Comput Appl Biosci* **13**(1): 23-8.
- Salzberg, S. L., M. Pertea, et al. (1999). "Interpolated Markov models for eukaryotic gene finding." *Genomics* **59**(1): 24-31.

- Scorilas, A. (2002). "Polyadenylate polymerase (PAP) and 3' end pre-mRNA processing: function, assays, and association with disease." Crit Rev Clin Lab Sci **39**(3): 193-224.
- Sheets, M. D., S. C. Ogg, et al. (1990). "Point mutations in AAUAAA and the poly (A) addition site: effects on the accuracy and efficiency of cleavage and polyadenylation in vitro." Nucleic Acids Res **18**(19): 5799-805.
- Simmer, F., C. Moorman, et al. (2003). "Genome-wide RNAi of *C. elegans* using the hypersensitive rrf-3 strain reveals novel gene functions." PLoS Biol **1**(1): E12.
- Stabenau, A., G. McVicker, et al. (2004). "The Ensembl core software libraries." Genome Res **14**(5): 929-33.
- Stajich, J. E., D. Block, et al. (2002). "The Bioperl toolkit: Perl modules for the life sciences." Genome Res **12**(10): 1611-8.
- Stanke, M. and S. Waack (2003). "Gene prediction with a hidden Markov model and a new intron submodel." Bioinformatics **19 Suppl 2**: II215-II225.
- Stein, L. D., Z. Bao, et al. (2003). "The genome sequence of *Caenorhabditis briggsae*: a platform for comparative genomics." PLoS Biol **1**(2): E45.
- Tabaska, J. E. and M. Q. Zhang (1999). "Detection of polyadenylation signals in human DNA sequences." Gene **231**(1-2): 77-86.
- Thastrom, A., L. M. Bingham, et al. (2004). "Nucleosomal locations of dominant DNA sequence motifs for histone-DNA interactions and nucleosome positioning." J Mol Biol **338**(4): 695-709.
- Tian, B., J. Hu, et al. (2005). "A large-scale analysis of mRNA polyadenylation of human and mouse genes." Nucleic Acids Res **33**(1): 201-12.
- Timchenko, L. T. (1999). "Myotonic dystrophy: the role of RNA CUG triplet repeats." Am J Hum Genet **64**(2): 360-4.
- Touchon, M., A. Arneodo, et al. (2004). "Transcription-coupled and splicing-coupled strand asymmetries in eukaryotic genomes." Nucleic Acids Res **32**(17): 4969-78.
- Vagner, S., C. Vagner, et al. (2000). "The carboxyl terminus of vertebrate poly(A) polymerase interacts with U2AF 65 to couple 3'-end processing and splicing." Genes Dev **14**(4): 403-13.
- Viterbi, A. (1967). "Error bounds for convolutional codes and an asymptotically optimum decoding algorithm." IEEE Transactions on Information Theory: 260-269.
- Wahle, E., A. Lustig, et al. (1993). "Mammalian poly(A)-binding protein II. Physical properties and binding to polynucleotides." J Biol Chem **268**(4): 2937-45.
- Walhout, A. J. and M. Vidal (2001). "Protein interaction maps for model organisms." Nat Rev Mol Cell Biol **2**(1): 55-62.
- Wickens, M., D. S. Bernstein, et al. (2002). "A PUF family portrait: 3'UTR regulation as a way of life." Trends Genet **18**(3): 150-7.
- Wilkie, G. S., K. S. Dickson, et al. (2003). "Regulation of mRNA translation by 5'- and 3'-UTR-binding factors." Trends Biochem Sci **28**(4): 182-8.
- Xu, N., C. Y. Chen, et al. (1997). "Modulation of the fate of cytoplasmic mRNA by AU-rich elements: key sequence features controlling mRNA deadenylation and decay." Mol Cell Biol **17**(8): 4611-21.
- Zhang, H., J. Hu, et al. (2005). "PolyA_DB: a database for mammalian mRNA polyadenylation." Nucleic Acids Res Database Issue: D116-20.
- Zhang, M. Q. (2002). "Computational prediction of eukaryotic protein-coding genes." Nat Rev Genet **3**(9): 698-709.

Zhao, J., L. Hyman, et al. (1999). "Formation of mRNA 3' ends in eukaryotes: mechanism, regulation, and interrelationships with other steps in mRNA synthesis." Microbiol Mol Biol Rev **63**(2): 405-45.