

## Jan Mrázek's publications

1. **Mrázek, J.** (2009) Finding sequence motifs in prokaryotic genomes – a brief practical guide for a microbiologist. *Briefings in Bioinformatics* 10, 525-536.
2. **Mrázek, J.** (2009) Phylogenetic signals in DNA composition: limitations and prospects. *Mol. Biol. Evol.* 26, 1163–1169.
3. **Mrázek, J.** and Summers, A.O. (2008) General characteristics of prokaryotic genomes. In: Xu, Y. and Gogarten, P. (Eds.) *Computational methods for understanding bacterial and archaeal genomes*. Imperial College Press, London, pp. 1-37
4. Guo, X. and **Mrázek, J.** (2008) Long simple sequence repeats in host-adapted pathogens localize near genes encoding antigens, housekeeping genes, and pseudogenes. *J. Mol. Evol.* 67, 497–509.
5. **Mrázek, J.**, Xie, S., Guo, X. and Srivastava, A. (2008) AIMIE: a web-based environment for detection and interpretation of significant sequence motifs in prokaryotic genomes. *Bioinformatics* 24, 1041-1048.
6. Schell, M.A., Ulrich, R.L., Ribot, W.J., Brueggemann, E.E., Hines, H.B., Chen, D., Lipscomb, L., Kim, H.S., **Mrázek, J.**, Nierman, W.C. and DeShazer, D. (2007) Type VI secretion is a major virulence determinant in *Burkholderia mallei*. *Mol. Microbiol.* 64, 1466-1485.
7. **Mrázek, J.**, Guo, X. and Shah, A. (2007) Simple sequence repeats in prokaryotic genomes. *Proc. Natl. Acad. Sci. USA* 104, 8472-8477.
8. **Mrázek, J.** and Karlin, S. (2007) Distinctive features of large complex virus genomes and proteomes. *Proc. Natl. Acad. Sci. USA*, 104, 5127-5132.
9. **Mrázek, J.** and Xie, S. (2006) Pattern Locator: A new tool for finding local sequence patterns in genomic DNA sequences. *Bioinformatics* 22, 3099-3100.
10. Badger, J.H., Hoover, T.M., Brun, Y.V., Weiner, R.M., Laub, M.T., Alexandre, G., **Mrázek, J.**, Ren, Q., Paulsen, I.T., Nelson, K.E., Khouri, H.M., Radune, D., Sosa, J., Dodson, R.J., Sullivan, S.A., Rosovitz, M.J., Madupu, R., Brinkac, L.M., Durkin, A.S., Daugherty, S.C., Kothari, S.P., Giglio, M.G., Zhou, L., Haft, D.H., Selengut, J.D., Davidsen, T.M., Yang, Q., Zafar, N., and Ward, N.L. (2006) Comparative genomic evidence for a close relationship between the dimorphic prosthecate bacteria *Hyphomonas neptunium* and *Caulobacter crescentus*. *J. Bacteriol.*, 188, 6841–6850.
11. Karlin, S., Brocchieri, L., **Mrázek, J.** and Kaiser, D. (2006) Distinguishing features of  $\delta$ -proteobacterial genomes. *Proc. Natl. Acad. Sci. USA*, 103, 11352-11357.
12. **Mrázek, J.** (2006) Analysis of distribution indicates diverse functions of simple sequence repeats in *Mycoplasma* genomes. *Mol. Biol. Evol.* 23, 1370-1385.
13. **Mrázek, J.**, Spormann, A. M. and Karlin, S. (2006) Genomic Comparisons among  $\gamma$ -proteobacteria. *Env. Microbiol.* 8, 273-288.
14. Karlin, S., **Mrázek, J.**, Ma, J. and Brocchieri, L. (2005) Predicted highly expressed genes in archaeal genomes. *Proc. Natl. Acad. Sci. USA* 102, 7303-7308.
15. Karlin, S., Brocchieri, L., Campbell, A., Cyert, M. and **Mrázek, J.** (2005) Genomic and proteomic comparisons between bacterial and archaeal genomes and related comparisons with the yeast and fly genomes. *Proc. Natl. Acad. Sci. USA* 102, 7309-7314

16. Karlin, S., Theriot, J. and **Mrázek, J.** (2004) Comparative Analysis of Gene Expression among Low G+C Gram-Positive Genomes. *Proc. Natl. Acad. Sci. USA* 101, 6182-6187.
17. Karlin, S., **Mrázek, J.** and Gentles, A.J. (2003) Genome comparisons and analysis. *Curr. Opin. Struct. Biol.* 13, 344-352.
18. Karlin, S., Barnett, M.J., Campbell, A.M., Fisher, R.F. and **Mrázek, J.** (2003) Predicting gene expression levels from codon biases in  $\alpha$ -proteobacterial genomes. *Proc. Natl. Acad. Sci. USA* 100, 7313-7318.
19. **Mrázek, J.** (2002) New technology may reveal mechanisms of radiation resistance in *Deinococcus radiodurans*. *Proc. Natl. Acad. Sci. USA* 99, 10943-10944.
20. **Mrázek, J.**, Gaynon, L.H. and Karlin, S. (2002) Frequent oligonucleotide motifs in genomes of three streptococci. *Nucleic Acids Res.* 30, 4216-4221.
21. Karlin, S., Brocchieri, L., Trent, J., Blaisdell, B.E. and **Mrázek, J.** (2002) Heterogeneity of genome and proteome content in bacteria, archaea, and eukaryotes. *Theor. Popul. Biol.* 61, 367-390.
22. Karlin, S., Brocchieri, L., Bergman, A., **Mrázek, J.** and Gentles, A.J. (2002) Amino acid runs in eukaryotic proteomes and disease associations. *Proc. Natl. Acad. Sci. USA* 99, 333-338.
23. Karlin, S., **Mrázek, J.**, Campbell, A. and Kaiser, D. (2001) Characterizations of highly expressed genes of four fast-growing bacteria. *J. Bacteriol.* 183, 5025-5040.
24. Karlin, S. and **Mrázek, J.** (2001) Predicted highly expressed and putative alien genes of *Deinococcus radiodurans* and implications for resistance to ionizing radiation damage. *Proc. Natl. Acad. Sci. USA* 98, 5240-5245.
25. **Mrázek, J.**, Bhaya, D., Grossman, A. R., and Karlin, S. (2001) Highly expressed and alien genes of the *Synechocystis* genome. *Nucleic Acids Res.* 29, 1590-1601.
26. Karlin, S. and **Mrázek, J.** (2000) Predicted Highly expressed genes of diverse prokaryotic genomes. *J. Bacteriol.* 182, 5238-5250.
27. Karlin, S., Brocchieri, L., **Mrázek, J.**, Campbell, A. M., and Spormann, A. M. (1999) A chimeric prokaryotic ancestry of mitochondria and primitive eukaryotes. *Proc. Natl. Acad. Sci. USA* 96, 9190-9195.
28. Campbell, A., **Mrázek, J.**, and Karlin, S. (1999) Genome signature comparisons among prokaryote, plasmid, and mitochondrial DNA. *Proc. Natl. Acad. Sci. USA* 96, 9184-9189.
29. **Mrázek, J.** and Karlin, S. (1999) Detecting alien genes in bacterial genomes. *Ann. New York Acad. Sci.*, volume 870 "Molecular Strategies in Biological Evolution (eds., L.H. Caporale and W. Arber), pp. 314-329.
30. Karlin, S., **Mrázek, J.**, and Campbell, A. M. (1998) Codon usages in different gene classes of the *Escherichia coli* genome. *Mol. Microbiol.* 29, 1341-1355.
31. Karlin, S., Campbell, A. M., and **Mrázek, J.** (1998) Comparative DNA analysis across diverse genomes. *Annu. Rev. Genet.* 32, 185-225.
32. **Mrázek, J.** and Karlin, S. (1998) Strand compositional asymmetry in bacterial and large viral genomes. *Proc. Natl. Acad. Sci. USA* 95, 3720-3725.
33. Karlin, S. and **Mrázek, J.** (1997) Prokaryotic genome-wide comparisons and evolutionary implications. In: *Bacterial Genomes: Physical Structure and Analysis.* (eds. F.J. de Bruijn, G.M. Weinstock and J.R. Lupski), pp. 196-212, Chapman & Hall, New York.

34. Karlin, S. and **Mrázek, J.** (1997) Compositional differences within and between eukaryotic genomes. *Proc. Natl. Acad. Sci. USA* 94, 10227-10232.
35. Karlin, S., **Mrázek, J.**, and Campbell, A. M. (1997) Compositional biases of bacterial genomes and evolutionary implications. *J. Bacteriol.* 179, 3899-3913.
36. Karlin, S., **Mrázek, J.**, and Campbell, A. M. (1996) Frequent oligonucleotides and peptides of the *Haemophilus influenzae* genome. *Nucleic Acids Res.* 24, 4263-4272.
37. Karlin, S. and **Mrázek, J.** (1996) What drives codon choices in human genes? *J. Mol. Biol.* 262, 459-472.
38. **Mrázek, J.** and Karlin, S. (1996) A new significant recurrent dyad pairing in *Haemophilus influenzae*. *Trends Biochem. Sci.* 21, 201-202.
39. **Mrázek, J.** and Kypr, J. (1995) Middle range clustering of nucleotides in genomes. *Comput. Appl. Biosci.* 11, 195-199.
40. **Mrázek, J.** and Kypr, J. (1994) Biased distribution of adenine and thymine in gene nucleotide sequences. *J. Mol. Evol.* 39, 439-447.
41. **Mrázek, J.** and Kypr, J. (1993) UNIREP: A microcomputer program to find unique and repetitive nucleotide sequences in genomes. *Comput. Appl. Biosci.* 9, 355-360.
42. Balcarová, Z., **Mrázek, J.**, Kleinwächter, V. and Brabec, V. (1992) Cleavage by restriction enzymes of DNA modified with the antitumor drug cis-diamminedichloroplatinum(II). *Gen. Physiol. Biophys.* 11, 579-588.
43. **Mrázek, J.** and Španová, A. (1992) ANAGEL: A personal computer program for evaluating DNA fragment lengths from distances migrated in an agarose gel. *Comput. Appl. Biosci.* 8, 524.
44. **Mrázek, J.** and Kypr, J. (1992) Nucleotide composition of genes and hydrophobicity of the encoded proteins. *FEBS Lett.* 305, 163-165.
45. **Mrázek, J.** and Kypr, J. (1992) DNABIND: An interactive microcomputer program searching for nucleotide sequences which may code for conserved DNA-binding protein motifs. *Comput. Appl. Biosci.* 8, 401-404.
46. **Mrázek, J.** and Kypr, J. (1992) GLOBIC: A very fast microcomputer program for fingerprinting, characterization and comparison of long nucleotide sequences. *Comput. Appl. Biosci.* 8, 29-34.
47. **Mrázek, J.** and Kypr, J. (1991) Prediction of protein secondary structure. *Chemické listy* 85, 1203-1218 (in Czech).
48. Kypr, J., **Mrázek, J.** and Reich, J. (1989) Nucleotide composition bias and CpG dinucleotide content in the genomes of HIV and HTLV 1/2. *Biochim. Biophys. Acta* 1009, 280-282.
49. Kypr, J. and **Mrázek, J.** (1989) Reading frames of HIV genes. *J. Theor. Biol.* 141, 423-424.
50. **Mrázek, J.** and Kypr, J. (1988) Computer program Jamsek combining statistical and stereochemical rules for the prediction of protein secondary structure. *Comput. Appl. Biosci.* 4, 297-302.
51. Kypr, J. and **Mrázek, J.** (1987) Unusual codon usage of HIV. *Nature* 327, 20.
52. Kypr, J. and **Mrázek, J.** (1987) Occurrence of nucleotide triplets in genes and the secondary structure of the coded proteins. *Int. J. Biol. Macromol.* 9, 49-53.
53. Kypr, J. and **Mrázek, J.** (1987) Pseudogene in the genome of bacteriophage lambda? *Biochem. Biophys. Res. Commun.* 145, 330-335.

54. Kypr, J. and **Mrázek, J.** (1986) Lambda phage protein Nu 1 contains the conserved DNA binding fold of repressors. *J. Mol. Biol.* 191, 139-140.
55. Kypr, J. and **Mrázek, J.** (1985) Possible mechanism of the allosteric activation of cAMP receptor protein. *Biochem. Biophys. Res. Commun.* 131, 780-785.