References and further information for

Genomic phylogenies based on normalized BLASTP scores.

Robert L. Charlebois

NeuroGadgets Inc. and GenomeAtlantic

rlcharlebois@neurogadgets.com

The NeuroGadgets Inc. Bioinformatics Web Service (NGIBWS)

• http://www.neurogadgets.com/bws.php

• Charlebois, R.L., G.D.P. Clarke, R.G. Beiko & A. St. Jean. 2003. Characterization of species-specific genes using a flexible, web-based querying system. *FEMS Microbiol. Lett.* 225, 213-220.

Genomic phylogenies within NGIBWS

• Gophna, U., W.F. Doolittle & R.L. Charlebois. 2005. Weighted genome trees: refinements and applications. *J. Bacteriol.* 187, 1305-1316.

• Charlebois, R.L., R.G. Beiko & M.A. Ragan. 2004. Genome phylogenies. In: (R.P. Hirt & D.S. Horner, eds.) *Organelles, Genomes and Eukaryote Phylogeny: An Evolutionary Synthesis in the Age of Genomics*, CRC Press, Boca Raton FL, pp. 189-206.

• Clarke, G.D.P., R.G. Beiko, M.A. Ragan & R.L. Charlebois. 2002. Inferring genome trees using a filter to eliminate phylogenetically discordant sequences and a distance matrix based on mean normalized BLASTP scores. *J. Bacteriol.* 184, 2072-2080.

Other published studies utilizing NGIBWS

• Charlebois, R.L. & W.F. Doolittle. 2004. Computing prokaryotic gene ubiquity: rescuing the core from extinction. *Genome Res.* 14, 2469-2477.

• Gophna, U., R.L. Charlebois & W.F. Doolittle. 2004. Have archaeal genes contributed to bacterial virulence? *Trends Microbiol.* 12, 213-219.

• Ragan, M.A. & R.L. Charlebois. 2002. Distributional profiles of homologous open reading frames among bacterial phyla: implications for vertical and lateral transmission. *Int. J. Syst. Evol. Microbiol.* 52, 777-787.

EvolSimulator

• Beiko, R.G. & R.L. Charlebois. A simulation test bed for hypotheses of genome evolution. Submitted for publication.

Summary:
   A system and framework for evolutionary simulation experiments, featuring:
   —a nonstationary model of evolution, with flexible parameterization
   —speciation and extinction, constraining population size
   —sequence evolution, with genome-level rates and mutational biases, and gene- and residue-specific mutabilities
   —gene family evolution, with evolving retention probabilities and duplicabilities
   —lateral gene transfer by various scenarios, including random exchange, relations-biased, "niche"-constrained, gene complement-biased, and G+C-biased.

NGIBWS-style genomic phylogenies of EvolSimulator output

• Beiko, R.G., W.F. Doolittle & R.L. Charlebois. Work in progress.

Summary:
   EvolSimulator can produce output compatible with NGIBWS, thus enabling a comprehensive analysis of simulated genomic data. Genomic phylogeny is among such analyses, allowing one to further refine its methodology, and better understand its output, using controlled data sets.