

Date: Fri, 09 Mar 2001 06:35:56 -0800

From: Robert Jansen <rjansen@utxvms.cc.utexas.edu>

Subject: bot soc meetings

In-reply-to: <15016.19128.455887.995176@gargle.gargle.HOWL>

To: Bernard Moret <moret@cs.unm.edu>

Cc: "Linda A. Raubeson" <Linda.A.Raubeson@directory.reed.edu>,

Tandy Warnow <tandy@cs.utexas.edu>, "David A. Bader" <dbader@eece.unm.edu>,

Li-San Wang <lisan@cs.utexas.edu>, "Stacia K. Wyman" <stacia@cs.utexas.edu>

JANSEN, ROBERT K., 1* DAVID BADER2, BERNARD M. E. MORET2, LINDA A. RAUBESON3, LI-SAN WANG4, TANDY WARNOW4, and STACIA WYMAN4. 1Integrative Biology, University of Texas, Austin TX 78712, 2Computer Science, University of New Mexico, Albuquerque, NM 87131; 3Biological Sciences, Central Washington University, Ellensburg, WA 98926; Computer Science, University of Texas, Austin TX 78712. - New approaches for using gene order data in phylogeny reconstruction.

The rapid accumulation of whole genome sequences for a wide diversity of taxa is generating a huge amount of comparative data for biologists. The availability of whole genome sequences is providing a new set of molecular characters for phylogenetic reconstruction, which are especially useful for resolving deep branches of the tree of life. Changes in gene order are caused primarily by inversions, transpositions, and transversions. One of the major challenges for using genomic changes is the development of computational methods for handling these types of characters, especially in groups with large numbers of genes and highly rearranged genomes. We have been developing and testing a variety of methods for reconstructing phylogenies based on gene order data using both real and synthetic data. Two of the primary methods we have developed and tested are Maximum Parsimony on Binary Encodings (MPBE) and methods for correcting previous published distance measures (inversion and transposition distances). Our simulations show that all methods perform very well when the rates of change are low relative to the number of genes and that all methods perform poorly when rates of change are high relative to the number of genes. Furthermore, corrected distance measures greatly improve the accuracy of phylogenies. We have applied these new methods to a data set for the highly rearranged chloroplast genomes of the Campanulaceae. In this group, which generally has low rates of change relative to the number of genes, all methods recover congruent tree topologies. Gene order phylogenies for the Campanulaceae are considered accurate because they are largely congruent with trees generated from three chloroplast gene sequences (atpB, matK, and rbcL).

Robert K. Jansen

Chairman, Integrative Biology

Blake Centennial Professor

Patterson Laboratories

University of Texas at Austin

Austin TX 78712

Phone: (512)471-3878, (512) 232-5661

Fax: (512) 471-3878

Email: rjansen@utxvms.cc.utexas.edu

Web: <http://www.biosci.utexas.edu/IB/faculty/jansen.htm>