Genome analysis

ParIS Genome Rearrangement server

I. Miklós¹,*, P. Ittzés² and J. Hein³
¹Theoretical Biology and Ecology Modelling Group, Hungarian Academy of Sciences and Eötvös Loránd University, Pázmány Péter Sétány 1/c, H-1117 Budapest, Hungary, ²Collegium Budapest, Institute for Advanced Study, Szentháromság u. 2. H-1014 Budapest, Hungary and ³Genome Analysis and Bioinformatics Group, Oxford Centre for Gene Function, Department of Statistics, Oxford University, 1 South Parks Road, Oxford OX1 3TG, UK

Received on May 25, 2004; revised on June 29, 2004; accepted on July 7, 2004
Advance Access publication October 12, 2004

ABSTRACT

Summary: ParIS Genome Rearrangement is a web server for a Bayesian analysis of unichromosomal genome pairs. The underlying model allows inversions, transpositions and inverted transpositions. The server generates a Markov chain using a Partial Importance Sampler technique, and samples trajectories of mutations from this chain. The user can specify several marginalizations to the posterior: the posterior distribution of number of mutations needed to transform one genome into another, length distribution of mutations, number of mutations that have occurred at a given site. Both text and graphical outputs are available. We provide a limited server, a downloadable unlimited server that can be installed locally on any linux/Unix operating system, and a database of mitochondrial gene orders.

Contact: miklos@ramet.elte.hu

Genome rearrangement events consist of not only inversions and (for multichromosomal genomes) translocations but also transpositions and inverted transpositions. Unfortunately, parsimony algorithms which find the minimum number of mutations needed to transform one genome into another, length distribution of mutations, number of mutations that have occurred at a given site. Both text and graphical outputs are available. We provide a limited server, a downloadable unlimited server that can be installed locally on any linux/Unix operating system, and a database of mitochondrial gene orders.

The MCMC sampler and postscript generating algorithms are written in C. The web interface is written in php 4.3. The php code generates a Perl script, which runs in background

© The Author 2004. Published by Oxford University Press. All rights reserved. For Permissions, please email: journals.permissions@oupjournals.org
Fig. 1. The joint posterior distribution of number of inversions and transpositions rearranging fluke mitochondrial genomes of *Paragonimus westermani* and *Schistosoma japonicum*.

Fig. 2. The posterior length distribution of inversions and transpositions rearranging fluke mitochondrial genomes of *P. westermani* and *S. japonicum*.

Table 1. A shortest trajectory transforming mitochondrial genome *P. westermani* to that of *S. japonicum*

| 2 | 2 | 4 | 33 | 0 | 11 | 13 | 21 | 3 | 34 | 0 | 16 | 17 | 27 | 2 | 4 | 7 | 35 | 2 | 0 | 2 | 34 |

Mutations are separated by '|'. The first number indicates the type of mutation: 0, transposition; 1, inverted transposition where the right block is inverted; 2, inverted transposition where the left block is inverted; 3, inversion. Following numbers indicate the positions of breakpoints of that mutation.
Fig. 3. The expected number of mutations happened before and after genes in mitochondrial genomes of *P. westermani* and *S. japonicum*.

Fig. 4. The log-likelihood trace from the analysis of mitochondrial genomes of *P. westermani* and *S. japonicum.*
and invokes the compiled C codes. In the downloadable unlimited version, we have provided details on how to use the compiled C codes in command line, which helps the user to write short scripts driving long runs of investigations. The command line version works without having to install a web server.

ACKNOWLEDGEMENTS

We thank the Collegium Budapest for providing a web server for this project. Special thanks to Susan Hutchinson and Irmtraud M. Meyer for their help and to the anonymous referees for their useful comments. This work was funded by EPSRC grant HAMJW and MRC grant HAMKA. I.M. was also supported by a Békésy György postdoctoral fellowship.

REFERENCES