Inference of demographic parameters using importance sampling on coalescence history
Raphaël Leblois, M. de Iorio, R. Griffiths, François Rousset

To cite this version:
hal-02932346

HAL Id: hal-02932346
https://hal.inrae.fr/hal-02932346
Submitted on 7 Sep 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Inference of demographic parameters using importance sampling on coalescence history

Raphaël Leblois and François Rousset

Université Montpellier 2, CNRS, Institut des Sciences de l’Évolution, France
Unité Origine, Structure et Evolution de la Biodiversité, Museum National d’Histoire Naturelle, Paris, France

State of our art
We developed and evaluated the performance of maximum likelihood (ML) analysis of allele frequency data in a linear array of populations under isolation by distance (IBD) and in a model of isolation with migration (IM) between two populations.

Implementation in the software MIGRAINE
The Algorithms used in this study are implemented in MIGRAINE, a free program for likelihood analyses of genetic data, with a focus on spatially structured populations. The demographic models currently implemented in this program are:
- A simple model of isolation by distance in a linear habitat, as described in Rousset & Leblois (2007), which includes the infinite island model as a special case;
- An isolation with migration (IM) model with two populations. Note that this IM algorithm is still under development but will soon be available.

At the heart of the program is the importance sampling algorithm defined by de Iori & Griffiths (2004). Some approximate procedures (“PAC-likelihood”) are also available and have been shown to be at least as efficient than exact maximum likelihood for time constant models.

MIGRAINE is currently designed for allelic type data sets only, but a k-alleles mutation model is implemented for all demographic models and a strict stepwise mutation model (SMM) is implemented to some extent for models with one or two populations (not yet published). However, we plan to adapt MIGRAINE for sequence data in a “near” future.

Real data set analysis in a linear habitat: the damselfly data set from Watts et al. 2007 Mol. Ecol.
In this application, we found results congruent with our simulation tests (Fig. 5): estimation of the dispersal scale parameter seems to be much harder.

Few other demographic models such as single population with size fluctuations in time, n-population with constant migration and the possibility to consider multiple samples in time is almost ready to be tested. We are looking for students or post-docs to develop those models. Do not hesitate to contact us if you are potentially interested. Migraine will thus be regularly updated to consider more demographic and mutational models so that it can be used in a wider range of cases.

Check the Migraine web page regularly to find new model implementations and Download Migraine at:
http://kimura.univ-montp2.fr/~rousset/Migraine.htm

References


R. Leblois venue to SMBE 2008 was partially supported by a grant from the Centre National de la Recherche Scientifique (CNRS) Biodiversity : Endemism in New Caledonia: A phylogenetic and population-level study of its origins (MNHN, OSIB)