NEWS AND VIEWS

PERSPECTIVE

All models are wrong

MICHAEL J. HICKERSON*

*Department of Biology, City College of New York, 160 Convent Ave., MR 526, New York, NY 10031, USA, †The Graduate Center of the City University of New York, Subprogram in Ecology, Evolution and Behavior, 365 5th Ave, New York, NY 10016, USA

Abstract

As the field of phylogeography has continued to move in the model-based direction, researchers continue struggling to construct useful models for inference. These models must be both simple enough to be tractable yet contain enough of the complexity of the natural world to make meaningful inference. Beyond constructing such models for inference, researchers explore model space and test competing models with the data on hand, with the goal of improving the understanding of the natural world and the processes underlying natural biological communities. Approximate Bayesian computation (ABC) has increased in recent popularity as a tool for evaluating alternative historical demographic models given population genetic samples. As a thorough demonstration, Pelletier & Carstens (2014) use ABC to test 143 phylogeographic submodels given geographically widespread genetic samples from the salamander species *Plethodon idahoensis* (Carstens et al. 2004) and, in so doing, demonstrate how the results of the ABC model choice procedure are dependent on the model set one chooses to evaluate.

Keywords: amphibians, model choice, phylogeography, population genetics – empirical, population genetics – theoretical

Received 26 April 2014; revised 9 May 2014; accepted 12 May 2014

Historians have shied away from the quantitative methods that are commonplace in the historically focused scientific fields such as geology and evolutionary biology. Although there have been occasional attempts in using historical data to build and test alternative quantitative historical models (Turchin 2003, 2012), the resistance by most historians is usually based on arguing that the models used in such endeavours lack the necessary complexity and generality (Thomas 2014). As clearly outlined in this issue of Molecular Ecology by Pelletier & Carstens (2014), there has been a similar tension within the field of phylogeography and historical population genetic inference (Pelletier & Carstens 2014).

The field of phylogeography has continued to move in the model-based direction in recent years, and although there are various methods for statistical model comparison in phylogeography (Carstens & Knowles 2010; Carstens et al. 2013), ABC has gained recent popularity as a tool for evaluating alternative historical demographic models given population genetic data (Sunnäker et al. 2013). Using ABC to evaluate competing models has various hazards and comes with recommended precautions (Robert et al. 2011), and unsurprisingly, many if not most researchers have a healthy scepticism as these tools continue to mature. Rather than a ‘garbage in, garbage out’ problem of questionable model-based inferences coming from questionable data, researchers are often unsatisfied with the models themselves, due to the necessity of collapsing the natural world into a useful yet simplified representation. Indeed, as scientists, we are never ‘sure’ of anything (Feynman 1999), and ‘essentially all models are wrong, but some are useful’ (Box & Draper 1987) has become the norm in science as all models only approximate reality at best. Estimating parameters requires use of probabilistic models, and these estimates are only as good as the models that are used, yet researchers in phylogeography often find it challenging and/or unsatisfying to build and use models that are approximations of reality on the one hand while able to make useful inference on the other. Regardless of this resistance, phylogeography is undergoing an exciting growth phase, both with opportunities to collect genomic-scale data (McCormack et al. 2012) and the opportunity to test a bewildering array of complex demographic models (Excoffier et al. 2013).

Instead of using ABC to evaluate a diverse set of models, Pelletier & Carstens (2014) undertake a more focused investigation into the ‘innards’ of ABC using it to evaluate the likelihood of three of the simplest and most widely used coalescent-based models in phylogeography: the panmictic model, the two-population island model and the two-population divergence/isolation model (Fig. 1). Rather than using ABC with weighted model averaging to obtain the three corresponding posterior model probabilities while allowing for the handful of model parameters \((0, \tau, \gamma, M)\) to be estimated under each model conditioned on each model’s posterior probability, these three models are sliced up into 143 ‘submodels’ according to various parameter ranges. For example, models designated as 1020 and 1010 are both divergence/isolation models that only differ by the direction of migration. Following the strategy outlined in the paper, one would estimate the posterior...

© 2014 John Wiley & Sons Ltd
models are assigned diminishing levels of posterior probability.

This is an expected result, and implementation of the Bayes factor comparing the most probable model with all others can help account for this by weighting posterior probabilities by the numbers of prior models. Instead, the Bayes factor implemented in Pelletier & Carstens (2014) made pairwise comparisons between pairs of models having the highest and second highest posterior probabilities regardless of the number of models. This pairwise Bayes factor will always be more conservative except in cases when the posterior probabilities are equal for all models that are less probable than the most probable model. To illustrate given a case of five models with posterior probabilities of 0.7, 0.15, 0.15, 0.0 and 0.0, the pairwise Bayes factor would be ~4.7, whereas the “all-model” Bayes factor would be ~9.33, with the former being a more conservative value (i.e. moderately favouring the model with highest posterior probability instead of strongly favouring this model). Likewise, if the posterior probabilities are equal for all models that are less probable than the most probable model (i.e. posterior probabilities of 0.7, 0.075, 0.075, 0.075, 0.075), both Bayes factors would be ~9.33. In general, using the Bayes factor, incorporating all models is perhaps more consistent with the Bayesian approach of incorporating all uncertainty associated with the ABC model choice procedure. After all, one uses ABC model choice to evaluate all specified models rather than two of the total number of specified models, and this useful illustration by Pelletier & Carstens (2014) is a powerful reminder that the outcome of model comparison with ABC or other similar likelihood-based methods will always be dependent on the composition of the candidate model set.

The three general models explored by Pelletier & Carstens (2014) could be simple enough to allow the use of full-likelihood methods rather than ABC, and their exploration is by no means a recommendation for how such data should be analysed. Although ABC remains useful under complex models (Prado-Martinez et al. 2013; Robinson et al. 2013), using it for model choice comes with challenges and scaling up to genomic-scale data sets (i.e. next-generation sequencing; NGS) is computationally expensive (Sousa et al. 2011, but see Roux et al. 2013). There are non-ABC options that can accommodate and evaluate complex models given data sets that are typical using NGS including new methods exploiting the site frequency spectrum that can accommodate models of arbitrary complexity such as large numbers of populations, ghost populations, size change and asymmetrical migration (Excoffier et al. 2013). Another part of phylogeographic model space that is not easily explored is the spatially explicit component of genetic samples, yet there has been some exciting developments that may allow the nonspatial approximation to be relaxed (Barton et al. 2010; Brown & Knowles 2012; Kellerer et al. 2013).

As suggested by Pelletier & Carstens (2014), there might be reluctance to infer phylogeographic history using models due to the lack of prior knowledge about how one
might parameterize putative demographic histories, yet there are fortunately a number of model-based clustering algorithms that can help estimate how many populations the genetic samples of a species come from (Pritchard et al. 2000; Dawson & Belkhir 2001; Frichot et al. 2014), as well as ‘species delimitation’ methods if one is interested in assigning putative species names to populations having degrees of isolation that satisfy particular species models (Fujita et al. 2012). Provided that there are sufficient data, these delineation procedures can provide the model scaffolding for subsequent parameter estimation under a model or mixture of models that are in turn weighted by their relative posterior probabilities.

Although the outcome of model comparison with ABC or other similar likelihood-based methods will always be dependent on the composition of the model set, and parameter estimates will only be as good as the models that are used, model-based inference provides a number of benefits. Not only can one re-evaluate models against newly constructed models in subsequent studies, these evaluations are explicit and provide a standard measure of repeatability. As researchers in phylogeography realize that evaluation of panmictic populations and the assignment of individuals.

References


M.H. wrote this article.

doi: 10.1111/mec.12794