Dispatches

Evolutionary Ecology: Next Generation Inference

Oak forests support a rich ecology of fellow travellers, but how do these fare when the forests move during glacial cycles? The answers revealed by a new study are important for ecology, but being able to get answers at all highlights a turning point in evolutionary inference.

Stuart J.E. Baird

Inference is one of the fastest moving but least appreciated fields of research in evolutionary biology. When field leaders in host-parasite evolutionary ecology and inference team up, as they have to produce a paper [1] in this issue of *Current Biology*, we catch a glimpse of a new generation of inference tools. In reconstructing the assembly of an oak community over time and space using an explicit population genetic model, Stone *et al.* [1] show that range change over Pleistocene glacial cycles means equilibrium-based expectations are a poor guide to current ecological patterns. This may seem unsurprising, but equilibrium expectations have dominated evolutionary inference for decades because the alternative, more complicated models have been too difficult to analyse. That dominance is ending as next generation inference allows us to quantify our uncertainty regarding complex models.

Science explores systems in the laboratory to minimise as many sources of uncertainty as possible; results are then tied to the remaining uncertainty. Quantifying uncertainty is at the heart of the scientific endeavour, yet is often seen negatively. As Feynman [2] noted, scientists are criticized because we are (correctly) never *sure* of anything. Meanwhile, researchers, having studied a discipline, often resent needing to also learn some statistics in order to express belief in their findings. This attitude is especially inadequate in evolutionary biology because often our hypotheses are complicated and our systems do not fit inside a lab. We need *inference tools*, and we need to understand how they work so that we can ask answerable questions, design our data gathering effort focussed on relevant information, and gauge our belief in other people's work [3]. Hypothesis-driven research and

a sound inference process are the same thing.

A Kuhnian scientific revolution [4] has reshaped inference over the last three decades, with the expected high incidence of multiple independent discoveries of the same advances. There have been some memorable debates [5] but history has moved on and, by and large, null hypotheses and rules of thumb have been replaced by explicit models. The reasons are both fundamental and pragmatic. Information lies in surprises [6], and no one should be surprised when a simple null hypothetis is rejected because it poorly explains a complex pattern. While there is no such thing as a complicated null hypothesis, computers enable us to work with explicit models even when they become complex. As an example, gone are the days when the genetics of field samples from distant locations would be used either to demonstrate panmixis does not hold at all scales (unsurprising) or, in the absence of sufficient data, to reject this over-simple null hypothesis in order to justify further analysis assuming panmixis (unscientific). Now, model-based clustering algorithms [7–9] can be used to estimate how many units the field samples come from. What a relief. Examining past process, rather than current patterns, with explicit models of population splitting [10,11], we are beginning to study the demography of speciation, and with multi-taxon inference [1], co-evolutionary patterns.

Oaks have one of the richest associated insect faunas of any temperate forest tree. Graeme Stone's group has been interested in the co-evolution of gall wasps, which induce oak trees to build them homes (galls), and parasitoid wasps which, somewhat like cuckoos, take over these homes by inserting their own eggs into the gall. There are many taxa

of each, making this potentially a rich coevolution. The first step from interest to inference is to formulate an explicit model. It is useful here to mention the road not taken: the over-simple null hypotheses that could have easily been proposed and rejected to produce a sound but uninformative old-school publication. Oldgrowth oak forests give an impression of the stability and equilibrium (from a human perspective), one that inspired climax community theory in ecology. But science tells us man is *not* the measure of all things, and the pollen record shows that current forest assemblages are stochastic and ephemeral [12,13]. A null hypothesis of spatiotemporal equilibrium among oaks and their fellow travellers is easy to propose and would be easy to reject, given sufficient data. Such an analysis probably could have been published even though we would learn little from the effort, because nobody would be surprised.

Mike Hickerson [14] and Konrad Lohse [15] are among the new generation of inference specialists providing us with model-based alternatives to strawman null hypotheses. It should be obvious that this is hard work. The alternative to the spatiotemporal equilibrium null above is an explicit model that can approximate the spatial heterogeneity and temporal dynamics of the oak and their fellow travellers repeatedly colonising from, and retreating to, multiple glacial refugia over the cycles of the Pleistocene [16]. Explicit modelling of a single range expansion is already non-trivial [17,18], so how could the Stone/Lohse/Hickerson team [1] possibly hope to manage a model of *many* such events?

The answer is in the question. The authors [1] reduced the history of range changes and colonisations into a series of discrete instantaneous events affecting discrete homogeneous units. This may seem extreme, but all models are approximations and it is our job as scientists to gauge our belief in the inference process by understanding the nature of these approximations. In this case, we can imagine time-lapse

photography condensing the entire Pleistocene into a YouTube clip. Suppose it takes a long time for a wasp lineage to reach the Iberian glacial refuge from the Italian one, but once it gets there it prospers. In the time-lapse clip, Iberia will be free of the lineage one instant, and entirely colonised in the next. The discrete event model is a good approximation to the extent that the details skipped during such eye-blinks are irrelevant to the evolutionary outcome.

Rejecting the strawman null would allow us to state that the European Pleistocene history of oaks and their fellow travellers is not one of spatiotemporal equilibrium. Yawn. Next generation inference allows us not only to estimate the extent to which the fellow travellers have been associated over the dynamic history of the Pleistocene, but also, and of key importance, *to quantify our uncertainty regarding those estimates*. This is the big leap forward, because opinionated criticism of the conclusions can be replaced by informed scientific debate regarding the approximations in the model and the degree of uncertainty attached to the estimates. I judge that the multi-taxon co-estimation model used by Stone *et al.* [1] can reasonably approximate the evolutionary/ ecological history that resulted in the Oak wasp dataset. A match between model and data lends credibility to their results, which indicate associations between gall wasps and their parasitoids broke down during the dynamic range changes of the Pleistocene. Here is a surprise that captures the scope of what we learn from the study: the potentially 'rich coevolution' between these oak fellow travellers is, on examination, highly unlikely to have been elaborated during the last 2.6 million years because the close associations necessary for coevolution to occur were disrupted by the Pleistocene glacial cycles.

As more and more evolution researchers realise the inference models we use are an integral part of our work [19], rather than an inconvenient requirement for publication, the evolution of next generation inference will most likely outstrip the technological breakthroughs which have driven the pace of much recent research. Of all the sciences it appears ours may be the one that will advance most in the information age.

References

1. Stone, G.N., Lohse, K., Nicholls, J.A., Fuentes-Utrilla, P., Sinclair, F., Schönrogge, K., Csóka, G.r., Melika, G., Nieves-Aldrey, J.-L., Pujade-Villar, J., *et al.* (2012). Reconstructing community assembly in time and space reveals enemy escape in a western palaearctic insect community. Curr. Biol. *22*, 532–537.
2. Feynman, R.P. (1999). 'The Value of Science,' address to the National Academy of Sciences (Autumn 1955). In The Pleasure of Finding Things Out: The Best Short Works of Richard, P. Feynman and J. Robbins, eds. (Cambridge, Massachusetts: Perseus Publishing).
3. Fisher, R.A. (1925). Statistical Methods for Research Workers (Edinburgh: Oliver & Boyd).
4. Kuhn, T.S. (1962). The Structure of Scientific Revolutions (Chicago: University of Chicago Press).
5. Knowles, L.L. (2008). Why does a method that fails continue to be used? Evolution *62*, 2713–2717.
6. Edwards, A.W.F. (1972). Likelihood (Cambridge: Cambridge University Press).
7. Dawson, K.J., and Belkhir, K. (2001). A Bayesian approach to the identification of panmictic populations and the assignment of individuals. Genet. Res. *78*, 59–77.
8. Guillot, G., Mortier, F., and Estoup, A. (2005). Geneland: a computer package for landscape genetics. Mol. Ecol. Notes *5*, 712–716.
9. Pritchard, J.K., Stephens, M., and Donnelly, P. (2000). Inference of population structure using multilocus genotype data. Genetics *155*, 945–959.
10. Hey, J., and Nielsen, R. (2007). Integration within the Felsenstein equation for improved Markov chain Monte Carlo methods in population genetics. Proc. Natl. Acad. Sci. USA *104*, 2785–2790.
11. Wilson, I.J., Weale, M.E., and Balding, D.J. (2003). Inferences from DNA data: population histories, evolutionary processes, and forensic match probabilities. J. Roy. Stat. Soc. A *166*.
12. De Beaulieu, J.L., Miras, Y., Andrieu-Ponel, V., and Guiter, F. (2005). Vegetation dynamics in north-western Mediterranean regions: Instability of the Mediterranean bioclimate. Plant Biosystems *139*, 114–126.
13. Cook, J.E. (1996). Implications of modern successional theory for habitat typing: A review. For. Sci. *42*, 67–75.
14. Hickerson, M.J., Stahl, E., and Takebayashi, N. (2007). ms Bayes: pipeline for comparative phylogeographic inference using approximate Bayesian computation. BMC Bioinformatics *8*, 268.
15. Lohse, K., Harrison, R.J., and Barton, N.H. (2011). A general method for calculating likelihoods under the coalescent process. Genetics *189*, 977–U398.
16. Hewitt, G.M. (2004). Genetic consequences of climatic oscillations in the Quaternary. Philos. Trans. R. Soc. Lond. B Biol. Sci. *359*, 183–195.
17. Baird, S.J.E., and Santos, F. (2010). Monte Carlo integration over stepping stone models for spatial genetic inference using approximate Bayesian computation. Mol. Ecol. Resources *10*, 873–885.
18. Estoup, A., Baird, S.J.E., Ray, N., Currat, M., Cornuet, J.M., Santos, F., Beaumont, M.A., and Excoffier, L. (2010). Combining genetic, historical and geographical data to reconstruct the dynamics of bioinvasions: application to the cane toad Bufo marinus. Mol. Ecol. Resources *10*, 886–901.
19. Philippe, H., and Roure, B. (2011). Difficult phylogenetic questions: more data, maybe; better methods, certainly. BMC Biol. *9*, 91.

Research Centre in Biodiversity and Genetic Resources (CIBIO), University of Porto, Campus Agrário de Vairão, Rua Armando Quintas 7, 4485-661 Vairão, Portugal; and French National Institute for Agricultural Research (INRA), 147 rue de l'Université, 75338 Paris Cedex 07, France.
E-mail: Stuart.J.E.Baird@gmail.com

Pheromones: Fish Fear Factor

Fish, like many other animals, panic when another individual is injured. Now, the chemical nature of a substance that mediates this reaction has been uncovered.

Marcus C. Stensmyr[1] and Florian Maderspacher[2]

Karl Ritter von Frisch — the Austro-German biologist best known for his work on the senses and communication of honeybees — was passionate about biology. So passionate that he once allegedly knocked a blueberry cake out of his wife's hands because she had dared to interrupt a scientific discussion at the coffee table; so passionate that he even used his summer holidays spent at the idyllic Wolfgangsee to do experiments (Figure 1). In the early 1930s, von Frisch was interested in the question of whether fish can hear. He had conditioned a swarm of minnows, a common small freshwater fish, to sounds and wanted to label the conditioned fish: "*To label another member of the shoal, I one day caught a minnow, severed its sympathicus nerve with a thin knife near the end of the tail, which causes a darkening of the skin caudal of the incision site, and set it free. There something unexpected*