

UC Merced

Frontiers of Biogeography

Title

Emerging patterns and emerging challenges of comparative phylogeography

Permalink

<https://escholarship.org/uc/item/2dm9n6kr>

Journal

Frontiers of Biogeography, 6(4)

Author

Horne, John B.

Publication Date

2014

DOI

10.21425/F5FBG24486

Copyright Information

Copyright 2014 by the author(s). This work is made available under the terms of a Creative Commons Attribution License, available at

<https://creativecommons.org/licenses/by/4.0/>

commentary

Emerging patterns and emerging challenges of comparative phylogeography

Biogeographic studies commonly amass distribution datasets for hundreds of species in an attempt to describe biogeographic patterns and their underlying processes (e.g., Keith et al. 2013). In contrast, phylogeographic studies are almost always limited to a small number of species and, while able to detect patterns at a population level, lack the replicative power of biogeography to describe those patterns in terms of relevant processes. This is particularly disadvantageous in phylogeographic studies of marine species, because the ocean contains few conspicuous geographic features that might explain population breaks (Horne 2014).

Large, comparative phylogeographic datasets have long been the daydream of molecular ecologists but have previously not been possible because the data didn't exist and rigorous phylogeographic analyses do not scale well with large datasets (Andrew et al. 2013). For instance, coalescence-based phylogeographic model fitting is able to accurately assess population patterns amid stochastic signals and other noise in the data (Beaumont et al. 2010) but is computationally intensive and can become overwhelmingly time consuming as datasets grow larger and modelled scenarios become more complex.

Even while data remain limited and coalescence analyses continue to be computationally burdensome, some research groups are expanding the boundaries of comparative phylogeographic and population genetic research. Specifically, a recent paper, published earlier this year, by Selkoe and colleagues, jointly analyzed population genetic patterns in 35 coral reef-associated species across the main and northwest Hawaiian islands. By biogeography standards this is a tiny dataset, but for a phylogeographic study 35 species is impressive.

Selkoe et al. (2014) did not include coalescence-based analyses, instead relying on

multivariate analyses (principal components analysis and redundancy analysis) of population genetic summary statistics, k-means clusters, and ecological variables to assess four types of population structuring: long-term panmixia across the 2400-km-long archipelago, chaotic genetic heterogeneity, isolation-by-distance (IBD), and regional genetic structure. The multivariate approach used was based, in part, on another comparative study of 27 high-alpine plants (Miermans et al. 2011). Since we can probably expect more studies like these to arise in the future, it seems timely to comment on these comparative methods, their drawbacks and their merits.

First, the entire approach rests upon a foundation of summary statistics (e.g., F_{ST}), which suffer from well-known shortcomings. For example, several of the datasets in Selkoe et al. did not have enough genetic polymorphism to statistically reject a null hypothesis of panmixia using gene frequencies. At the other extreme, some of the datasets had too much polymorphism to reject panmixia, because if each individual has a unique genetic variant the maximum attainable F_{ST} value is 0. Issues of polymorphism might be addressed with additional sampling, which could uncover additional genetic signal (e.g., Horne et al. 2013). Alternatively, one could collect more loci. Regardless, insufficient data is the underlying problem.

Ample amounts of genetic variation are the *sine qua non* of this method; a problem, considering that researchers are generally forced to rely on available datasets from past surveys, many of which were kept small by the costs of Sanger DNA sequencing. Furthermore, species that are too sparsely sampled, or contain too few individuals at each location, may not capture enough genetic variation. Selkoe et al. rightly point out that comparative studies demand substantial rigor in sampling coverage

and consistency. Although, some species may have unexplainable rarity or absence on local scales.

Collating different datasets, with different sampling strategies and intensities into a single study is no small task; even a single phylogeographic data set can be influenced by numerous factors. Reconciling the noise in 35 datasets is not simple. It is here that the benefit of the multivariate analyses really shines. In Selkoe et al. (2014) and Miermans et al. (2011), multivariate analysis accounted for 11 and 17% of the variation in the data, respectively. Though much of the variation in the data remains unexplained, past studies suggest that weak genetic patterns are often biologically significant (Eble et al. 2009, Horne et al. 2013).

Among the more interesting results of Selkoe et al. (2014) was that only four species exhibited IBD – a pertinent result considering the nearly linear array of habitat. IBD was also associated with invertebrates with shallow depth requirements. The idea that depth range is correlated with dispersal ability is an old notion in biogeography (Brown et al. 1996) but one that has gained recent momentum. Keith et al. (2013) found that for every 10 m increase in depth range, coral taxa were 27% more likely to straddle faunal breaks. Von der Heyden et al. (2013) found that the vertical positioning of shoreline fishes was a significant predictor of population structure.

Curiously, all of the species with chaotic genetic heterogeneity were non-endemics and habitat generalists, while most habitat specialists and endemics exhibited regional population structure. It is difficult to explain these results but the emergence of such patterns showcases the utility of the multivariate comparative approach. Yet, perhaps the most valuable result of Selkoe et al. (2014) is the unexplainable variation in the data. Some of this unexplained variation is probably attributable to random historical events, i.e. differences in the location of first colonization, or the num-

ber of independent colonizations of each species. Other variation might be attributable to genetic drift and other stochastic processes.

A few summary statistics not included by Selkoe et al. are a neutrality test (e.g., Tajima's D) and mismatch distribution values τ and θ . Like all summary statistics, these suffer from specific shortcomings, and the latter two are for mtDNA only, but future studies might consider investigating noise in the data caused by selection and the relative impacts of expansion age and effective population size, as these are related to genetic drift. In the meantime, we are left to ponder whether stochastic signals account for most of the variation in spatial genetic patterns of communities, or if there are better ecological frameworks to consider.

Large comparative phylogeographic studies now offer us the opportunity to re-examine the relationship between ecology and genetic patterns and question our understanding of population-level processes. Notwithstanding present methodological challenges, the biggest limitation of comparative phylogeography is the availability of data. The data of Selkoe et al. (2014), while one of the largest datasets of its kind, represented only a minute fraction of the total Hawaiian marine biodiversity. Therefore, in spite of the fact that basic phylogeographic studies are not as publishable as they used to be, basic phylogeographic data, of all types, are needed now more than ever.

John B. Horne

Center of Marine Sciences,
University of Algarve, Faro, Portugal 8005-139
john.horne@gmail.com

References

- Andrew, R.L., Bernatchez, L., Bonin, A., et al. (2013) A road map for molecular ecology. *Molecular Ecology*, 22, 2605–2626.
- Brown, J.H., Stevens, G.C. & Kaufman, D.M. (1996) The geographic range: size, shape, boundaries and internal structure. *Annual Review of Ecology and Systematics*, 27, 597–623.

- Beaumont, M.A., Nielsen, R., Christian, R., et al. (2010) In defense of model-based inference in phylogeography. *Molecular Ecology*, 19, 436–446.
- Eble, J.A., Toonen, R.J. & Bowen, B.W. (2009) Endemism and dispersal: comparative phylogeography of three surgeonfishes across the Hawaiian archipelago. *Marine Biology*, 156, 689–698.
- Horne, J.B. (2014) Thinking outside the barrier: neutral and adaptive divergence in Indo-Pacific coral reef faunas. *Evolutionary Ecology*, 28, 991–1002.
- Horne, J.B., Momigliano, P., van Herwerden, L. & Newman, S.J. (2013) Murky waters: searching for structure in genetically depauperate blue threadfin populations of Western Australia. *Fisheries Research*, 146, 1–6.
- Keith, S.A., Baird, A.H., Hughes, T.P., Madin, J.S. & Connolly, S.R. (2013) Faunal breaks and species composition of Indo-Pacific corals: the role of plate tectonics, environment and habitat distribution. *Proceedings of the Royal Society of London Series B*, 280, 20130818.
- Meirmans, P.G., Goudet, J. & Gaggiotti, O.E. (2011) Ecology and life history affect different aspects of the population structure of 27 high-alpine plants. *Molecular Ecology*, 20, 3144–3155.
- Selkoe, K.A., Gaggiotti, O.E., TOBO Laboratory, Bowen, B.W. & Toonen, R.J. (2014) Emerging patterns of population genetic structure for a coral reef community. *Molecular Ecology*, 23, 3064–3079.
- von der Heyden, S., Gildenhuis, E., Bernardi, G. & Bowie, R.C.K. (2013) Fine-scale biogeography: tidal elevation strongly affects population genetic structure and demographic history in intertidal fishes. *Frontiers of Biogeography*, 5, 29–38.

Submitted: 10 November 2014

First decision: 16 November 2014

Accepted: 01 December 2014

Edited by Michael N Dawson and Chris BurrIDGE

Your participation in **frontiers of biogeography** is encouraged. Please send us your articles, comments and/or reviews. We are also open to suggestions on content and/or structure.

Please check <http://www.biogeography.org/html/fb.html> for more information, or contact us at ibs@mncn.csic.es and frontiersofbiogeography@gmail.com.