NEWS AND VIEWS

COMMENT

The value of DNA sequence data for studying landscape genetics

ANDREW J. BOHONAK* and AMY G.
VANDERGAST†

*Department of Biology, San Diego State University, 5500
Campanile Drive, San Diego, CA 92182-4614, USA, †U.S.
Geological Survey, Western Ecological Research Center, San
Diego Field Station. 4165 Spruance Road, Suite 101, San Diego,
CA 92101, USA

Abstract

In a recent Opinion article in *Molecular Ecology*, Wang
(2010) emphasizes the fact that current patterns of genetic
differentiation among populations reflect processes that
have acted over temporal scales ranging from contempo-
rary to ancient. He draws a sharp distinction between the
fields of phylogeography (as the study of historical pro-
cesses) and landscape genetics (which he restricts to very
recent processes). Wang characterizes DNA sequence data
as being inappropriate for the study of contemporary
population processes and further states that studies
which only include mitochondrial DNA or chloroplast
DNA data cannot be considered part of landscape genet-
ics. In this response, we clarify the generally accepted
view that DNA sequence data can be analysed with
methods that separate contemporary and historical pro-
cesses. To illustrate this point, we summarize the study
of Vandergast *et al.* (2007), which Wang mischaracterizes
as being confused in terms of temporal scale. Although
additional focus should be placed on the important issue
of correct data interpretation, we disagree strongly with
the implication that contemporary and historic processes
cannot be separated in the analyses of DNA sequence
data.

Keywords: cp/mtDNA, DNA sequence data, landscape
genetics, nonequilibrium conditions, phylogeography

*Received 27 October 2010; revision received 7 December 2010;
accepted 27 December 2010*

The Opinion article of Wang (2010) focuses on the defini-
tion and application of phylogeography and landscape
genetics. Because different genetic metrics and molecular
markers integrate over different lengths of time, it is

Correspondence: Andrew J. Bohonak, Fax: 619 594 5676;
E-mail: bohonak@sciences.sdsu.edu

important that researchers interpret their results in terms
of the correct temporal scale. Wang advocates restriction of
the term 'landscape genetics' to specific types of genetic
data, analyses and temporal scales and states that the 'gap
between phylogeography and landscape genetics cannot be
thoroughly investigated with current methods.' He twice
cites one of our studies (Vandergast *et al.* 2007) as being
confused on these points. In this response, we clarify some
technical issues from his overview of the field, and elabo-
rate on the approaches taken in our study.

Wang (2010) defines landscape genetics as being focused
on contemporary processes, clearly distinct from the study
of historic processes in the field of phylogeography. Within
this framework, he has several critiques of intraspecific
genetic studies. He first concerns the types of markers that
are used to study contemporary processes. After acknowl-
edging that population genetic structure in neutral markers
is the product of mutation, drift and gene flow, Wang cau-
tions that chloroplast DNA and mitochondrial DNA
(cp/mtDNA) evolve 'too slowly to be useful for inferring
most recent and ongoing microevolutionary processes.
Although some noncoding cp/mtDNA regions evolve at
relatively faster rates than coding regions, these still experi-
ence considerably lower substitution rates than typical mi-
crosatellite loci.' The substitution rate is the rate at which
new mutations become fixed, which is simply equal to the
mutation rate. Thus, Wang implicitly focuses on analytical
methods that require populations to be fixed for unique
mutations. His concerns about substitution (or mutation)
rate for organellular sequence data would also apply to
nuclear sequence data.

Both cp/mtDNA and nDNA sequence data are analysed
with a variety of methods. The fixation of new mutations in
isolated populations is almost never a criterion of interest in
the studies of contemporary processes, regardless of the
marker used. More commonly, analyses focus on allele fre-
quency differences and the genealogical relationships (evo-
lutionary distances) between alleles. Mutation, drift and
population-specific patterns of gene flow collectively influ-
ence allele frequencies over a range of temporal scales that
always begin with the most recent generation. In fact,
cp/mtDNA markers may be better suited for some studies
of contemporary evolutionary processes than nuclear mark-
ers. If sufficiently variable sequence data are available, drift
will be faster in organellular genomes because of their lower
effective population size. It is possible that Wang's critique
is very narrowly aimed at the interpretation of traditional
phylogeographic methods (such as labelling gene trees with
geographic locations), and if so, his point is valid.

Wang's second concern is that at least some methods for
analysing intraspecific genetic data are being misinter-
preted in terms of the wrong temporal scale. This concern

has traditionally been expressed in terms of the amount of time a particular metric or marker requires to reach equilibrium. The recognition and study of nonequilibrium conditions has been a prominent part of the literature over the past three decades (e.g., Slatkin 1985; Wade & McCauley 1988; Boileau *et al.* 1992; Bossart & Pashley Prowell 1998; Bohonak & Roderick 2001; Excoffier *et al.* 2009; Allen *et al.* 2010). It can be summarized briefly as follows:

1 Genealogy-based analyses that require new, unique mutations to reach detectable levels or even fixation in local populations must necessarily focus on the oldest evolutionary processes in the gene genealogy. Thus, phylogeographic analyses (Avise *et al.* 1987) often aim to infer historic isolating events from the accumulation of fixed differences among lineages over time. Similarly, cladistic gene flow estimates (Slatkin & Maddison 1989) and coalescent-based models (e.g., Beerli & Felsenstein 1999; Hey & Nielsen 2004) implicitly assume that parameters such as gene flow and drift have been constant across the entire gene genealogy. Thus, contemporary changes in gene flow rates or patterns are not readily inferred from these methods.

2 Analyses that use frequency-based similarity or distance measures (AMOVA, Excoffier *et al.* 1992; e.g., IBD, Slatkin 1993) may reach drift–gene flow equilibrium on shorter timescales than coalescent-based analyses, typically requiring at least tens and possibly thousands of generation. The specific amount of time required depends on whether gene flow (or effective population size) has increased or decreased and the magnitude of the change. Frequency-based genetic distance measures can provide information even prior to a drift–gene flow equilibrium when comparative approaches are employed (e.g., Bohonak & Roderick 2001; Keyghobadi 2007 and references therein). For example, the levels of genetic differentiation can be compared in recently fragmented and contiguous habitats.

3 Clustering algorithms commonly applied to microsatellite data (e.g., STRUCTURE, Pritchard *et al.* 2000; BAPS, Corander *et al.* 2003) are used to define contemporary gene pool boundaries. These analyses utilize linkage disequilibrium across loci, which is statistically detectable for only a few generations after a unique genotype immigrates.

Clearly, these three categories of analysis overlap in terms of the time periods for which they provide useful information. Nonetheless, Wang (2010) states that the temporal gradient between phylogeographic studies and landscape genetics 'is an oversimplification' and that there is 'a gap between phylogeography and landscape genetics which cannot be thoroughly investigated with current methods.' In contrast, we believe that the sharp dichotomy is an oversimplification, because many analyses bridge that gap by separating historical from current population processes. Perhaps Wang is most concerned about the use of terminology: 'studies should adhere to the definitions of phylogeography and landscape genetics… particularly in

regard to their explicit temporal distinctions, and employ the correct terminology so that readers clearly understand the validity of the inferences made in those studies.' Although we agree that population geneticists must understand the temporal bounds of their analyses, we disagree as to how often published studies make this mistake or are in turn misinterpreted by other researchers in the field.

Our interpretation of Wang's opinion piece is that he wishes to restrict the term 'phylogeography' to describe DNA sequence data analysed with genealogical approaches (category 1 above) and the term 'landscape genetics' to microsatellite data analysed with allele frequency-based statistics or clustering algorithms (categories 2–3). Coalescent models are mentioned in his review, but their placement into either phylogenetics or landscape genetics is ambiguous. Wang also marginalizes the dominant role that frequency-based analyses of DNA sequences have played in developing our understanding of both recent and historic evolutionary processes since the 1980s. Although Wang cites Manel *et al.* (2003) as providing a model definition for landscape genetics that should not be 'diluted', he overlooks the fact that Manel *et al.*'s review discussed multiple studies that only used mtDNA. In contrast to Wang's assertion, we believe that mtDNA data can be used to make accurate inferences about contemporary processes such as spatial patterns of gene flow. We illustrate this point using our recent study that he cites as an example of misapplication.

Vandergast *et al.* (2007) studied the populations of the flightless Jerusalem cricket *Stenopelmatus 'mahogani'* in southern California using mtDNA sequence data. The sampling sites were located throughout a landscape that included both large and small habitat fragments surrounding the Los Angeles basin. We analysed these data using four approaches:

1 As is common, the regional sampling locations were labelled on a Bayesian gene genealogy and on a parsimony network to highlight deep historical isolation. The genealogy was dated using a molecular clock. Regions were not reciprocally monophyletic, and individual sampling locations often shared alleles within a region.

2 To separate recent population isolation from that which occurred prior to human influence, we generated a spatial (GIS-based) model of habitat fragmentation by marine inundation during the Holocene/Pleistocene. The effects of distance, ancient and current habitat fragmentation were analysed using partial Mantel tests and a novel regression approach that is described in the study. The results clearly demonstrate that both contemporary and historic landscape features influence current patterns of mtDNA differentiation. This approach has been advocated in several recent review articles in *Molecular Ecology* (e.g., Anderson *et al.* 2010).

3 Correlations between genetic diversity and contiguous habitat area were considered in terms of both prehistoric and current fragment size.

4 One large historic fragment in our study has been recently bisected by a highway. Genetic divergence

across the highway is higher than for comparable distances on the same side of the highway. Computer simulations showed that increased mtDNA divergence across the highway matches theoretical expectations for complete cessation of gene flow on a timescale corresponding to the highway's construction.

In summary, Vandergast *et al.* (2007) interpreted mtDNA diversity using multiple approaches that include both phylogeography and landscape genetics. Beyond our explanations of the timescale relevant for each analysis, we actually validated our interpretations with a dated gene genealogy, historic and contemporary GIS models and computer simulations. Although the study could be criticized for using only one genetic marker (a point with which we do not disagree), Wang's primary concern is instead that the article's title includes the term 'landscape genetics'. His implication that we are confused and incorrectly interpreted our analyses seems unwarranted.

Wang (2010) cites several other studies in support of his general argument, but we could not find fault with their data interpretation either. Swart *et al.* (2009) and Measey & Tolley (in press) both used mtDNA data to study ancient evolutionary events in terms of landscape-level processes. Each listed the keyword phrase 'landscape genetics', and both interpreted their data in terms of the appropriate temporal scale. Wang also cited Koscinski *et al.* (2009) because their study relied on mtDNA but included the keyword phrase 'landscape genetics'. Like Vandergast *et al.* (2007), Koscinski *et al.* (2009) analysed their data in terms of both historic processes and current landscape variables. We believe that 'landscape genetics' is not a misleading study descriptor in this case.

In summary, we welcome the additional attention that Wang's review has brought to the appropriate timescales (time to equilibrium) for different types of analyses and molecular markers. However, we disagree strongly with the implication that contemporary processes cannot be analysed using DNA sequence data. We also suggest that it is not fruitful to restrict the term 'landscape genetics' to studies that only focus on contemporary processes. A major goal of landscape genetics is to test for the effects of recent habitat change, and this can be carried out by comparing current and past landscape features using approaches like those described in Vandergast *et al.* (2007), regardless of marker type.

References

Allen MR, Thum RA, Caceres CE (2010) Does local adaptation to resources explain genetic differentiation among *Daphnia* populations? *Molecular Ecology*, **19**, 3076–3087.

Anderson CD, Epperson BK, Fortin MJ *et al.* (2010) Considering spatial and temporal scale in landscape-genetic studies of gene flow. *Molecular Ecology*, **19**, 3565–3575.

Avise JC, Arnold J, Ball RM *et al.* (1987) Intraspecific phylogeography: the mitochondrial bridge between population genetics and systematics. *Annual Review of Ecology and Systematics*, **18**, 489–522.

Beerli P, Felsenstein J (1999) Maximum-likelihood estimation of migration rates and effective population numbers in two populations using a coalescent approach. *Genetics*, **152**, 763–773.

Bohonak AJ, Roderick GK (2001) Dispersal of invertebrates among temporary ponds: are genetic estimates accurate? *Israel Journal of Zoology*, **47**, 367–386.

Boileau MG, Hebert PDN, Schwartz SS (1992) Non-equilibrium gene frequency divergence: persistent founder effects in natural populations. *Journal of Evolutionary Biology*, **5**, 25–39.

Bossart JL, Pashley Prowell D (1998) Genetic estimates of population structure and gene flow: limitations, lessons and new directions. *Trends in Ecology and Evolution*, **13**, 202–206.

Corander J, Waldmann P, Sillanpaa MJ (2003) Bayesian analysis of genetic differentiation between populations. *Genetics*, **163**, 367–374.

Excoffier L, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics*, **131**, 479–491.

Excoffier L, Foll M, Petit RJ (2009) Genetic consequences of range expansions. *Annual Review of Ecology Evolution and Systematics*, **40**, 481–501.

Hey J, Nielsen R (2004) Multilocus methods for estimating population sizes, migration rates and divergence time, with applications to the divergence of *Drosophila pseudoobscura* and *D. persimilis*. *Genetics*, **167**, 747–760.

Keyghobadi N (2007) The genetic implications of habitat fragmentation for animals. *Canadian Journal of Zoology*, **85**, 1049–1064.

Koscinski D, Yates AG, Handford P, Lougheed SC (2009) Effects of landscape and history on diversification of a montane, stream-breeding amphibian. *Journal of Biogeography*, **36**, 255–265.

Manel S, Schwartz MK, Luikart G, Taberlet P (2003) Landscape genetics: combining landscape ecology and population genetics. *Trends in Ecology and Evolution*, **18**, 189–197.

Measey GJ, Tolley KA (2011) Investigating the cause of the disjunct distribution of *Amietophrynus pantherinus*, the endangered South African western leopard toad. *Conservation Genetics*, **12**, 61–70.

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics*, **155**, 945–959.

Slatkin M (1985) Gene flow in natural populations. *Annual Review of Ecology and Systematics*, **16**, 393–430.

Slatkin M (1993) Isolation by distance in equilibrium and non-equilibrium populations. *Evolution*, **47**, 264–279.

Slatkin M, Maddison WP (1989) A cladistic measure of gene flow inferred from the phylogenies of alleles. *Genetics*, **123**, 603–614.

Swart BL, Tolley KA, Matthee CA (2009) Climate change drives speciation in the southern rock agama (*Agama atra*) in the Cape Floristic Region, South Africa. *Journal of Biogeography*, **36**, 78–87.

Vandergast AG, Bohonak AJ, Weissman DB, Fisher RN (2007) Understanding the genetic effects of recent habitat fragmentation in the context of evolutionary history: phylogeography and landscape genetics of a southern California endemic Jerusalem cricket (Orthoptera: Stenopelmatidae: *Stenopelmatus*). *Molecular Ecology*, **16**, 977–992.

Wade MJ, McCauley DE (1988) Extinction and recolonization: their effects on the genetic differentiation of local populations. *Evolution*, **42**, 995–1005.

Wang IJ (2010) Recognizing the temporal distinctions between landscape genetics and phylogeography. *Molecular Ecology*, **19**, 2605–2608.