Simulating the effects of migration rates on Neolithic range expansion

Neha J. Angal and Christopher R. Tillquist

Anthropology, University of Louisville

Abstract
In this simulation study, we investigated how migration rates, fecundity, and carrying capacity may have influenced the development of diversity during a range expansion. Using NEMO, an open source simulation environment for population genetics, we simulated range expansions, which typically generate a diversity cline. Migration occurred as a 2D-stepping-stone model with migration rates of 0.2 and 0.3. We modeled ten lattice lattices, tracked average heterozygosity for 1024 demes, and sampled from the first generation where the lattice filled and the final generation. Sampling at the time all demes are filled, showed independent effects of migration rates, fecundity, and carrying capacity on lattice-wide heterozygosity.

Background
Continental colonization events can result in differential spatial diversity due to serial founder effects and allele surfing during range expansion events. Establishing which factors result in the unique patterns during an expansion may provide insights into colonization processes. These processes and resulting distributions have been investigated for several decades in Europe using complex models and simulation studies, with a particular emphasis placed upon determining if Europe was peopled in the late Paleolithic or the Neolithic (Arenas et al. 2012, Arenas et al. 2013, Klijnstra et al., 2016, Bandi, et al., 1996, Bartończyk et al., 1995, Wegman et al. 2016). However, many of these studies relied on small lattices that are less representative of continental space, and utilized programs that are not publicly available. Hence results that are non-repeatable, and relied upon parameters that may not be realistic for modeling human populations.

Objectives and Hypotheses
The goal of this study was to determine what distributions of diversity develop in the context of a range expansion, which parameters drive those distributions, and how long signals can persist. We expected that migration rate would have the strongest apparent effect upon distributions of diversity.

Materials & Methods
Forward simulations were performed using NEMO (Guillaume et al. 2006). Result files were processed with gene scripts, PDGSpiter (Lacher and Escorih, 2012), and analyzed using Arlequin 3.5. Files were post-processed with gene scripts and Vim. Heterozygosities were calculated using the Hygen package (Athey and Franca, 2015) and plotted in R using the heatmaps module of gplots (Warnes et al., 2015). Preserving and plotting of Arlequin results files were completed using R version 3.2.4 (Very Secure Directions, R Core Development Team).

Experimental Design
Two sets of simulations were run for experimental and control scripts: a set of one stage and a set of two stage models. Experimental scripts modeled range expansions across a mostly empty lattice, and control scripts modeled instantaneous-fitted lattices. Both lattices were 32 by 32 for a total of 1024 demes. Ten lattices were replaced for each simulation, and all ten were randomly polymorphic at the start, such that allele frequencies were set at 0.5. One stage models varied migration rates (either 0.2 or 0.3), mean fecundity (either 3 or 4), and carrying capacity (either 200 or 500). The second stage models were the same as above, but included an extra stage with replicate runs to confirm the results with the first stage lasting for 250 generations and the second stage lasting for 500 generations, for a total of 750 generations. Twenty replicate runs were completed for each simulation. For each experimental simulation, results files for the first generation at which the lattice was filled (determined by each patch containing at least 25 individuals) and the final generation were obtained; corresponding files for control simulations at the same generation were also obtained. Parameter choices were selected based on previous simulations in the literature.

Results
Across all combinations of controls versus simulated range expansions, only heatmaps of the first-generation-to-fin differences. While at the final generations there were no differences between control and experimental simulations, there was variability in the degree of diversity with experimental simulation showing lower diversity. With regard to the one stage models, all differed between control and experimental simulations except for the simulations where migration was 0.2, mean fecundity was 3, and carrying capacity was 500.

The results of varied combinations of parameters differed, and in all cases lower diversity was observed in the experimental simulations unless the migration rate was 0.3, where the mean fecundity was 4, and carrying capacity was 200. In these, both the control and experimental simulations had low diversity at generation 500 and appeared very similar to one another. As a whole, the one stage models documented the influence of each parameter separately on diversity, holding other variables constant; higher carrying capacity resulted in higher diversity; higher migration resulted in higher diversity; and higher fecundity resulted in higher diversity, with the most marked difference observed when both carrying capacity and fecundity were high.

Summary
Simulations of range expansion models using variable migration rates, fecundity, and carrying capacity all resulted in loss of diversity. Parameter choice played a strong role in the overall results, such that a lower migration rate (m=0.2), lower fecundity (f=3), and lower carrying capacity (cc=200), resulted in lower diversity in demes outside of the origin. Generally, these patterns followed the theoretical expectations of a model of serial founder effects. Results from our simulations are in accord with empirical data, supporting a strong role of drift in generating human population structure and the complexity of continental colonization events.

Future planned simulations will work to tease out the interactions between the three parameters in the context of selection.

Acknowledgments
This work was conducted in part using the resources of the University of Louisville’sResearch Computing Group and the Computational Cluster. We would like to acknowledge Dr. Forrest Slevin (Department of Geography and Geosciences, University of Louisville) for his invaluable assistance with data management, and Dr. Jeremy Biggs (CIAM, Collection and Information Access Management), CIAM, at the Kentucky Center for the Advancement of Science and Technology, University of Louisville, for his invaluable assistance with data management. In R, we used the library pbdR (http://cran.r-project.org/web/packages/pbdR/index.html).

Supporting information and data are available on the University of Louisville’s Research Computing Group’s Research Computing website (http://researchcomputing.uofl.edu).

References