

# From there to here: Biological distance in prehistoric West-Central Illinois

Susan R. Frankenberg and Lyle W. Konigsberg  
Department of Anthropology  
University of Illinois at Champaign-Urbana

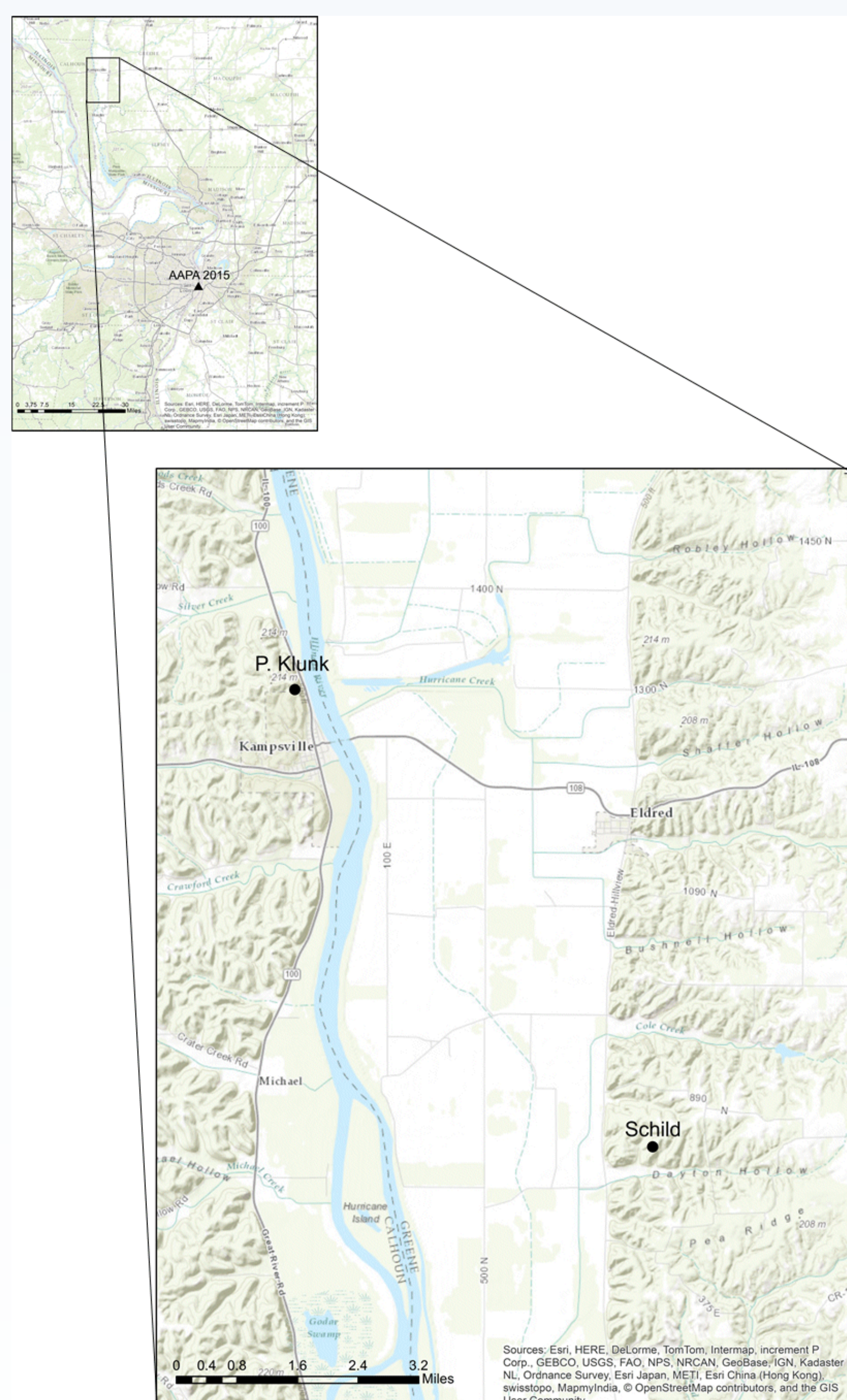


## Introduction

Biological distance analysis for prehistoric West-Central Illinois has shifted from a descriptive focus to one including multivariate tests of statistical hypotheses. However, connections of these hypotheses to population genetic parameters is tenuous at best. Even with the advent of ancient DNA analysis, there has been a tendency to avoid model-bound methods to estimate population genetic parameters. We show using aDNA haplogroup data from Bolnick and Smith (2007) and Raff (2008) how model-bound Bayesian analyses can be constructed in place of classical statistical hypothesis testing. Haplogroup data does not provide sufficient information for estimation, but haplotype data shows greater promise. It remains doubtful whether phenotypic data can provide enough information to estimate population genetic parameters.

## The Samples

The haplogroup data are from Bolnick (2005) for the Pete Klunk Site (Middle Woodland component) and Raff (2008) for the Schild Site (Mississippian component).



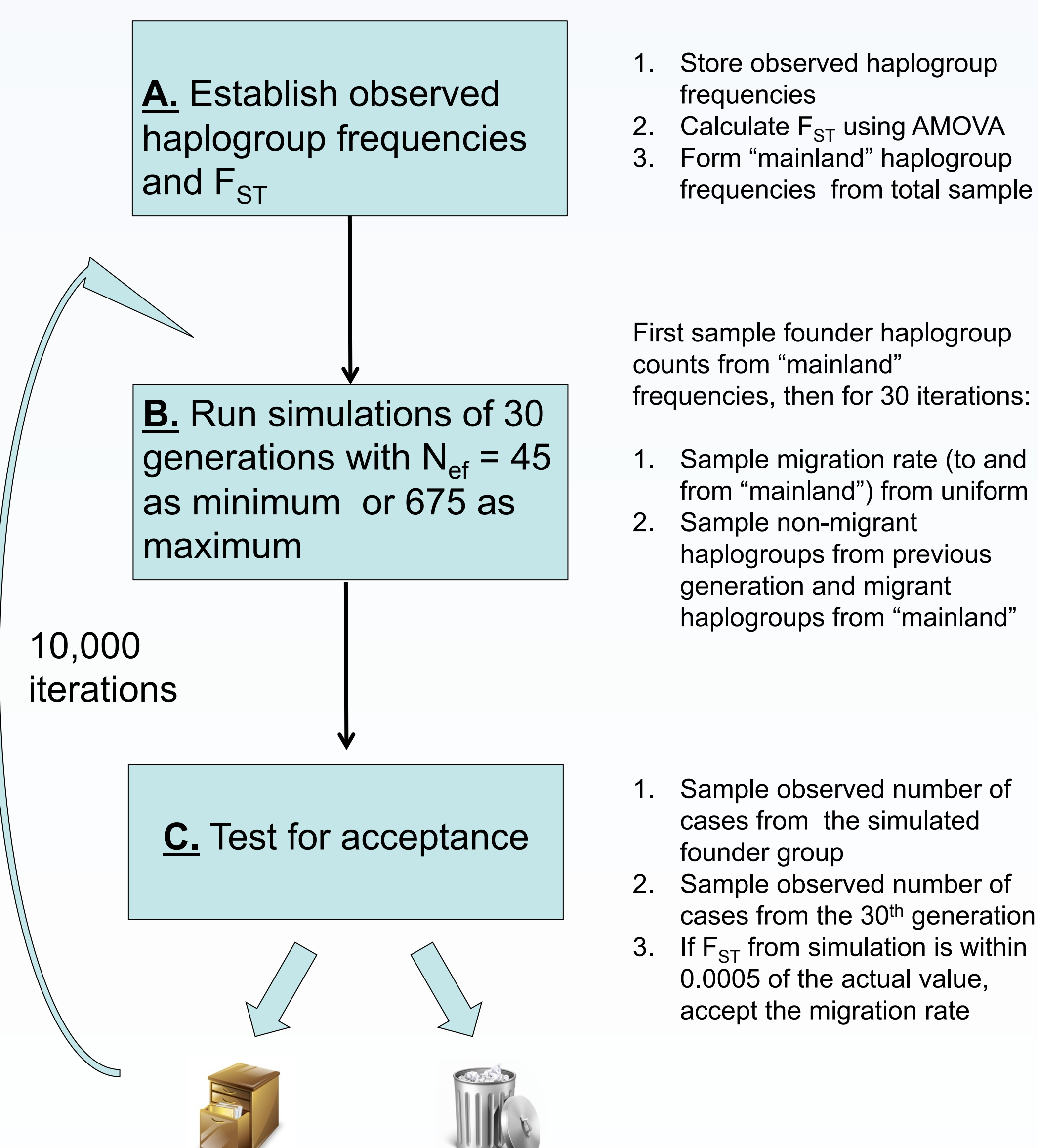
## Approximate Bayesian Computation (ABC)

Cabana et al. (2008) proposed an elaborate simulation method to test whether the amount of haplogroup differentiation (genetic distance) between an ancestral and a presumed descendant sample exceeded that expected from genetic drift and short-range (adjacent deme) migration alone. Raff applied this test in her dissertation using Klunk and Schild haplogroup frequencies. Under reasonable effective population size estimates and short-range migration she concludes “that genetic drift is unlikely to account for these discontinuities.” She wrote that her data provided an “interpretation of the genetic discontinuity as resulting from gene flow, rather than genetic drift.”

We instead use ABC to estimate the long-range migration rate (between any given deme and the “mainland” in Wright’s island model). The data and simulation scheme are:

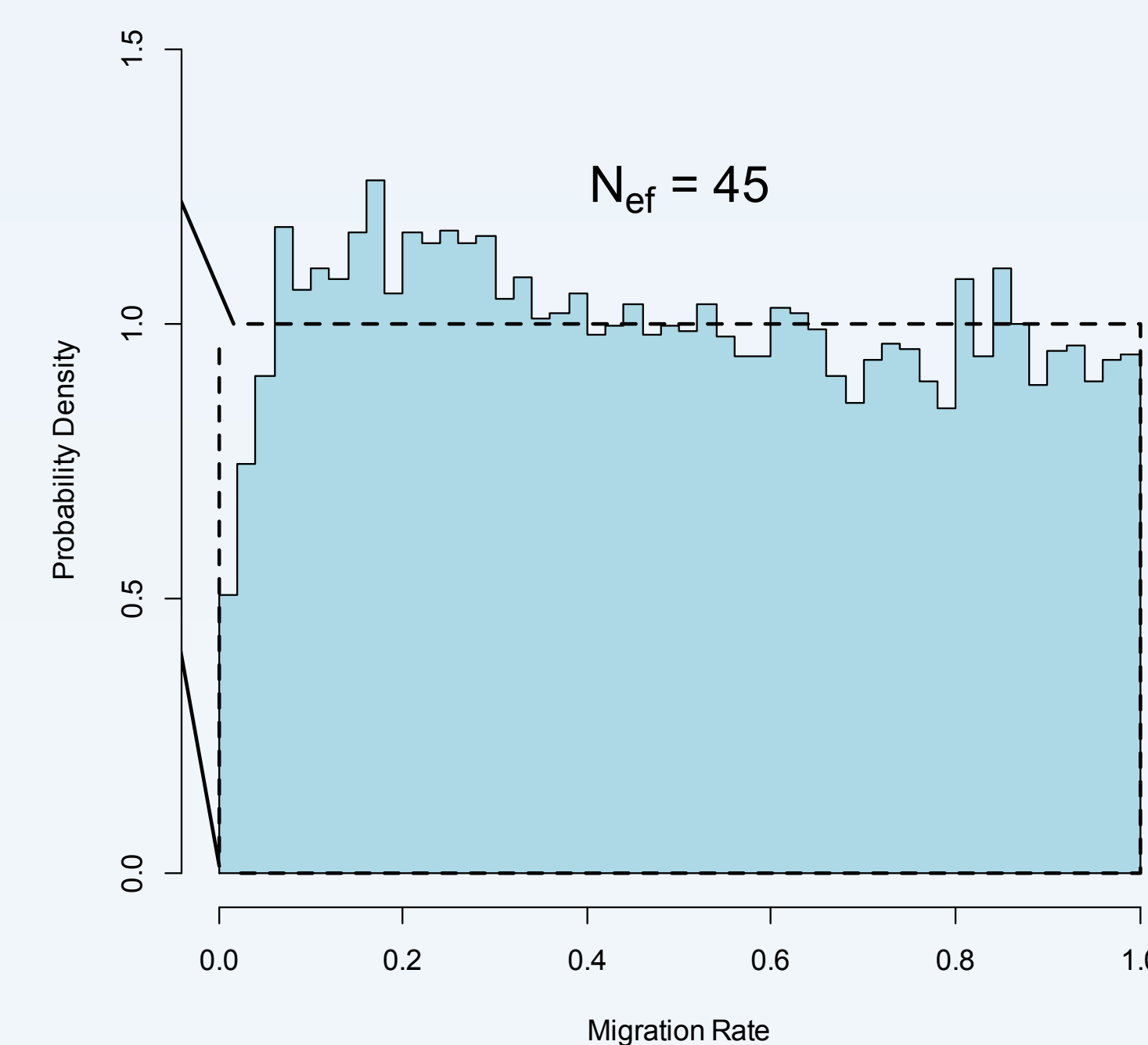
### Haplogroup frequencies

Site	A	B	C	D	X	Total
Pete Klunk	9	5	19	5	1	39
Schild	18	6	11	4	8	47

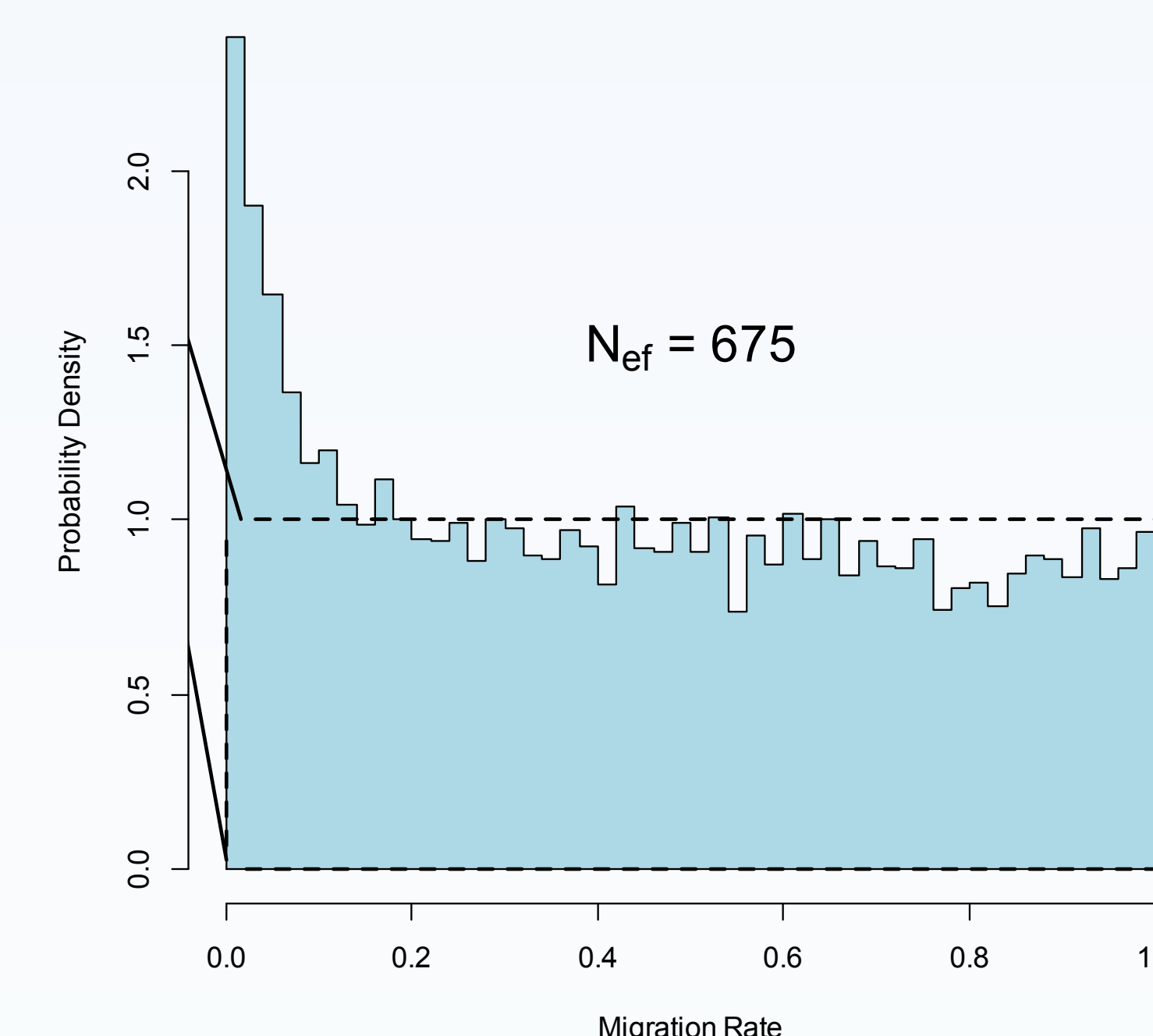


## Results

ABC with the observed haplogroup counts for 39 individuals from Pete Klunk and 47 individuals from Schild, 30 generations, and an effective (female) population size of 45 (equivalent to a minimum census size of 300) produces a distribution of migration rates that is little different from the prior uniform distribution. The 95% empirical highest posterior density for the migration rate is from 0.06 to 0.99.



At the maximum effective population size of 675 (representing a census size of 4,500 individuals) the posterior density for the migration rate does differ from the prior uniform distribution. But there is still considerable uncertainty represented in the posterior distribution, with the 95% highest posterior density (HPD) spanning from 0.00 to 0.94.



The problems in forming useful HPDs stem from sample size issues. At a ten-fold increase in the sample size and with an effective population size of 675, the HPD is from 0.00 to 0.03. With a 100-fold increase in sample size and with an effective population size of 45, the 95% HPD is only reduced from 0.06 – 0.99 (original sample size) to 0.00 – 0.88.

## Discussion

It is clear that haplogroup data alone are insufficient for estimating relevant population genetic parameters. The aDNA field has shifted to using haplotype data. Indeed, Reynold’s et al. (2015) have used haplotype data to test for continuity between Schild Late Woodland and Mississippian. Their analysis does not directly address the estimation of population genetic parameters, but instead maintains a hypothesis testing format.

Biological distance analysis using skeletal and/or dental data from West-Central Illinois has remained “stuck” in a hypothesis testing mode. For example, applications of the Relethford-Blangero model (Steadman 1998, 2001) have looked at the expected negative correlation between phenotypic variation and distance to the centroid and have identified “outliers” that may represent demes with greater or lesser gene flow than expected. But this is not the same thing as estimating effective population sizes or migration rates. Similarly, Konigsberg (1990) has used Mantel matrix correlation tests to examine the relationship between temporal, spatial, and biological distances for West-Central Illinois.

If the field is to move from “there to here” then we may have to admit that phenotypic traits are not informative enough to allow estimation of population genetic parameters. Even with haplogroup or haplotypic data from single loci, sample sizes may be inadequate to address questions of interest.

## References

- Bolnick DA, and Smith G. 2007. Migration and social structure among the Hopewell. *Amer Antiq* 72:627-644.
- Konigsberg LW. 1990. Analysis of prehistoric biological variation under a model of isolation by geographic and temporal distance. *Hum Biol* 62:49-70.
- Raff JA. 2008. An ancient DNA perspective on the prehistory of the Lower Illinois Valley. Doctoral dissertation, Indiana University.
- Reynolds AW, Raff JA, Bolnick DA, Cook DC, and Kaestle FA. 2015. Ancient DNA from the Schild site in Illinois: Implications for the Mississippian transition in the Lower Illinois River Valley. *Am J Phys Anthropol* 156:434-448.
- Steadman DW. 1998. The population shuffle in the central Illinois Valley: a diachronic model of Mississippian biocultural interactions. *Wld Archaeol* 30:306-326.
- Steadman DW. 2001. Mississippians in motion? A population genetic analysis of interregional gene flow in West-Central Illinois. *Am J Phys Anthropol* 114:61-73.