Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Genetics & Development

Integrating the signatures of demic expansion and archaic introgression in studies of human population genomics

Lauren Alpert Sugden[1,2] and Sohini Ramachandran[1,2]

Human population genomic studies have repeatedly observed a decrease in heterozygosity and an increase in linkage disequilibrium with geographic distance from Africa. While multiple demographic models can generate these patterns, many studies invoke the serial founder effect model, in which populations expand from a single origin and each new population's founders represent a subset of genetic variation in the previous population. The model assumes no admixture with archaic hominins, however, recent studies have identified loci in *Homo sapiens* bearing signatures of archaic introgression. These results appear to contradict the validity of analyses invoking the serial founder effect model, but we show these two perspectives are compatible. We also propose using the serial founder effect model as a null model for determining the signature of archaic admixture in modern human genomes at different geographic and genomic scales.

Addresses
[1] Center for Computational Molecular Biology, Brown University, Providence, RI, USA
[2] Department of Ecology and Evolutionary Biology, Brown University, Providence, RI, USA

Corresponding author: Ramachandran, Sohini (sramachandran@brown.edu)

Current Opinion in Genetics & Development 2016, **41**:140–149

This review comes from a themed issue on **Genetics of human origin**

Edited by **Joshua M Akey** and **Anna Di Rienzo**

For a complete overview see the Issue and the Editorial

Available online 13th October 2016

http://dx.doi.org/10.1016/j.gde.2016.09.007

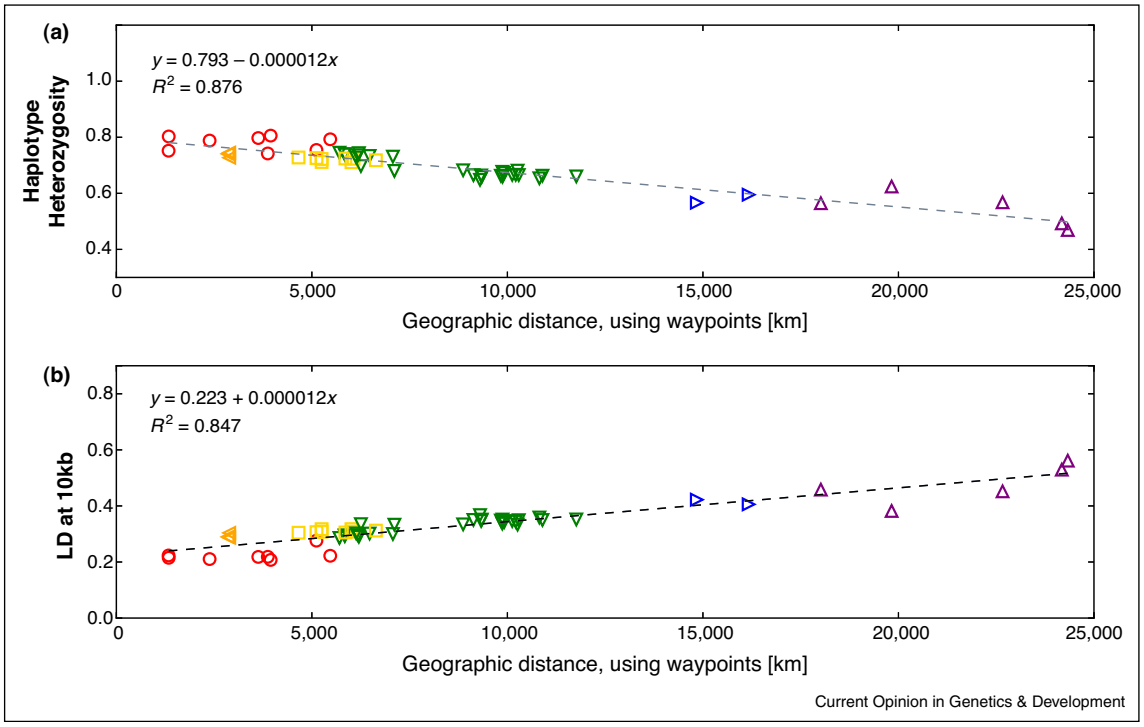
Introduction

Research in human evolution relies on multiple fields — such as archaeology, genetics, and linguistics — to give a history of *Homo sapiens* during the last hundred thousand years. In the last decade, investigations of worldwide human genomic variation based on multiple genetic marker types have observed three robust trends in summary statistics as a function of increasing geographic distance from Africa (Figure 1; see also DeGiorgio *et al.* [1]): a decrease in heterozygosity, an increase in linkage disequilibrium (LD), and an increase in the frequency of derived alleles [2–8]. Analyses of genomic datasets from human pathogens and parasites have also observed a reduction in heterozygosity as a function of distance from Africa [9–11]. Further, studies of anthropometric [12,13], economic [14], cultural [15], and linguistic data [16] have reported that population-level diversity in various traits follows this same pattern.

These studies each interpreted the observed reduction in trait diversity out of Africa as reflecting the population and geographic expansion of modern humans in the last 45k–60k years by invoking a model for the great human expansion known as the 'serial founder effect model' [2,3,17–20]). In a serial founder effect model, a series of successive bottlenecks from a single origin of expansion produces a stepwise increase in genetic drift — and consequent decrease in genetic diversity — as a function of geographic distance from the origin. Although the ability to identify a unique model for human evolutionary history from genome-wide diversity patterns has generated debate across a range of disciplines (for example, [1,21–29,30•,31,32•]), the serial founder effect model has proven to be a useful framework for both understanding the dynamic nonequilibrium history of human populations and identifying origins of large-scale human population expansions [6,19,33–35]. However, emerging high-quality archaic genomes and genomic datasets sampling increasing numbers of individuals from diverse human populations have shed new light on the complex demographic events that characterize human evolutionary history. Thus, the time is ripe to examine where the assumptions of the serial founder effect model falter, and what we can learn from genomic regions and populations in which genomic diversity patterns diverge from those predicted by the model.

We focus here on one particular assumption: the serial founder effect model, as invoked by Prugnolle *et al.* [2] and Ramachandran *et al.* [3], implicitly assumes that admixture with archaic hominins did not play a substantial role in the evolutionary history of *H. sapiens* (see also DeGiorgio *et al.* [1], Pickrell and Reich [29]). Here, we review observed patterns of human genomic variation from large-scale studies in the last five years, examine this assumption closely in light of these empirical results, and propose a framework for disentangling the signatures of archaic introgression from those of modern human population interactions in worldwide human genomic datasets.

Figure 1



Observed genome-wide patterns of heterozygosity and linkage disequilibrium (LD) among worldwide human populations as functions of geographic distance from Addis Ababa, Africa (9 N, 38 E). Haplotype heterozygosity (panel **(a)**, calculated as in Conrad *et al.* [95]) and average LD at 10 kb (panel **(b)**, measured by r^2) were calculated across 640,034 SNPs genotyped in the Human Genome Diversity Panel [7]. Note error bars are smaller than symbols. Equations for fitted lines and each linear model's coefficient of determination are displayed within each panel. Symbols indicate geographic regions: red circles = Africa, orange left-pointing triangles = Middle East, yellow squares = Europe, green down-pointing triangles = Central/South and East Asia, blue right-pointing triangles = Oceania, purple up-pointing triangles = Americas; populations are assigned to regions as in Rosenberg *et al.* [38].

Human genomic variation reveals intercontinental clusters and intracontinental admixture

Analyses of human genomic data from globally distributed populations repeatedly find that multilocus genotypes from human populations produce genetic clusters largely corresponding to major geographic regions [36–41]. Within continents, inferred genetic clusters may identify geographically or culturally isolated populations, distinguish among various subsistence strategies, or reveal signatures of gene flow [8,42,43].

The existence of genetic clusters among worldwide human populations has been challenged for multiple reasons [44–46]. One point of contention regarding the 'clusteredness' of humans is particularly relevant to this review: Serre and Pääbo [47] questioned whether the identification of clusters from human multilocus genotype data is an artefact due to the sampling design of datasets such as the Human Genome Diversity Panel [4,5,37,48]. In response, Rosenberg *et al.* [38] showed that inferred clusters arise not from the geographic dispersion of sampled individuals, but rather are generated by small discontinuous jumps in genetic distance for population pairs on opposite sides of geographic barriers (Figure 2 and [44]).

The serial founder effect model assumes no gene flow between neighboring populations, and generates a positive correlation between genetic distance and geographic distance among population pairs [3]. As was observed by Rosenberg *et al.* [38] using microsatellite data, we find that the effect of a barrier that delineates a continental region — such as a large mountain range or a continental shelf — is to add to pairwise genetic distance between populations beyond the value predicted by geographic distance alone. In the case of HGDP SNP data analyzed in Figure 2, crossing a barrier adds 0.0703 to pairwise F_{ST}, the equivalent to traveling approximately 9764 km within a continental region.

Thus, Figure 2 strongly suggests that qualitative features of the serial founder effect model hold true for human evolutionary history: *modern human genetic diversity generally reflects a stepwise accumulation of genetic drift across geographic barriers, and intercontinental migration was not*