**Introduction**

Evolutionary clines—changes in the frequency of a genotype or phenotype over a geographical area (Huxley 1938, 1939)—have long served as model systems in evolutionary biology (Endler 1977). Clines arise and are maintained via the interplay of natural selection, genetic drift, and migration across an environmental gradient (Haldane 1948; Felsenstein 1975; Slatkin and Maruyama 1975; Endler 1977; Vasemägi 2006; Saccheri et al. 2008; Volis and Zhang 2010; Takahashi 2015). The multiple evolutionary mechanisms structuring clines has prompted their continued use by evolutionary biologists seeking to explore the relative contributions of non-adaptive and adaptive evolutionary processes in structuring patterns of genetic and phenotypic diversity and differentiation within natural populations.

While clines are often interpreted as strong evidence of adaptive evolution, non-adaptive processes (e.g. genetic drift and gene flow) may also generate covariance between morph frequencies and environmental gradients. For example, local genetic drift in combination with spatially restricted gene flow (i.e. isolation by distance, Wright 1943) can generate single-locus clines (Vasemägi 2006). Similarly, serial founder events can generate clines in additive quantitative traits (Colautti and Lau 2015) and phenotypic clines can arise through multiple introductions from a species’ native range during invasion (Keller et al. 2009). Disentangling the relative importance of stochastic and deterministic forces in the formation of clines is thus essential prior to invoking the role of selection in generating differentiation among populations.

A corollary of the formation of single-locus or additive trait clines via neutral processes is that clines in both directions should occur with equal frequency; assuming the same initial allele frequencies, alleles should be lost or fixed with equal probability under only the effects of genetic drift. Thus, the presence of multiple independent clines in the same direction (i.e. parallel clines) is strong evidence for the role of natural selection, as putative adaptations are unlikely to evolve repeatedly via stochastic forces (Samis et al. 2012). However, when traits have a non-additive genetic basis (e.g. epistasis), clines may occur more frequently in a particular direction because stochastic changes in allele frequencies at one locus may have a disproportionate effect on phenotype frequencies. For example, stochastic forces have caused the repeated loss of the Mendelian inherited, epistatically determined short-style (S) morph from tristylous populations of *Eichhornia paniculata* in North-eastern Brazil, Jamaica, and Cuba (Barrett et al. 1989, 2009, Husband and Barrett 1992a,b; Barrett 1993). The fact that drift can lead to directional changes in non-additive traits across multiple, independent populations means that the presence of parallel clines in such traits is insufficient evidence for the role of selection in generating adaptive differentiation. In such cases, selection should only be invoked upon observing more clines than would be expected under solely the effects of drift. Studies exploring trait differentiation across replicate, independent geographical gradients in selection and demography and that take explicit consideration of the genetic architecture of the trait(s) in question would provide the strongest test of the relative contribution of drift, selection, and gene flow in the formation of parallel clines.

Urbanization is one of the most widespread human disturbances on earth and it provides an excellent large-scale replicated system to study how adaptive and non-adaptive evolutionary processes contribute to the formation of parallel clines. Urban environments continue to expand across the globe and are a leading cause of species extinction (McKinney 2006; Seto et al. 2010), stimulating research into how they alter the evolutionary processes that enable species to persist under these drastically altered environmental conditions. The widespread fragmentation associated with urbanization has resulted in gradients in the strength of genetic drift and gene flow. For example, urban fragmentation has reduced genetic diversity among urban populations of white-footed mice (*Peromyscus leucopus*) in New York City (Munshi-South and Kharchenko 2010; Munshi-South 2012; Munshi-South et al. 2016) and fire salamanders (*Salamandra salamandra*) in Oviedo, Spain (Lourenço et al. 2017) due to increased genetic drift and reduced gene flow among urban populations. While the influence of urbanization on non-adaptive evolutionary processes is well-studied (Johnson and Munshi-South 2017), evidence is emerging that urban environments alter natural selection and species adaptation as well.

Urban environments are associated with variation in biotic and abiotic factors (McKinney 2006; Johnson et al. 2015), many of which can be potent agents of selection for many taxa. For example, selection associated with urbanization has driven parallel evolution of pollution resistance in Atlantic Killifish (*Fundulus heteroclitus*) (Reid et al. 2016) and greater thermal tolerance in Acorn ant (*Temnothorax curvispinosus*) populations across multiple cities in the United States (Diamond et al. 2017, Diamond et al., 2018 *this issue*). Recently, Thompson *et al.* (2016) identified parallel urban-rural clines in the frequency of plants producing hydrogen cyanide (i.e. cyanogenesis, HCN)—a potent antiherbivore defence—in populations of white clover (*Trifolium repens*) across multiple cities. They found that HCN defended genotypes were less frequent in urban populations in 3 of the 4 cities examined(Thompson et al. 2016). While the authors identified lower winter surface temperatures in urban populations as a putative selective agent structuring urban-rural cyanogenesis clines clines, they did not consider the alternative hypothesis that these clines could be caused by genetic drift. In this study, we use the cyanogenesis polymorphism in white clove as a model for exploring the conditions under which non-adaptive (e.g. genetic drift, gene flow) and adaptive (e.g. selection) processes can generate repeated clines in phenotypes with a non-additive genetic basis.

White clover is native to Eurasia but has been intentionally introduced into temperate regions worldwide due to its importance in agriculture (Burdon 1983; Kjærgaard 2003). White clover and the cyanogenic polymorphism have a long-history of study among evolutionary biologists as latitudinal, longitudinal and altitudinal clines in cyanogenesis have been reported in clover’s native (Daday 1954a,b; de Araújo 1976) and non-native (Daday 1958; Ganders 1990; Kooyers and Olsen 2012) ranges, with higher frequencies of HCN reported in populations from warmer (e.g. lower elevation and further south) (Daday 1965) and drier environments (Kooyers et al. 2014). HCN is controlled by two independently segregating Mendelian loci (*CYP79D15* and *Li*) and plants require a dominant allele at both loci to produce HCN. As such, this trait exhibits duplicate recessive epistasis (i.e. complimentary epistasis, Sackton and Hartl 2016) where recessivity at either locus (or both) results in individuals lacking HCN (i.e. acyanogenic, HCN–). This genetic architecture has important consequences for how stochastic changes in allele frequencies are expected to affect the frequency of HCN: the frequency of HCN can only increase via drift if the frequency of the dominant alleles at both underlying loci drift upward (Table 1). Any other combination of frequency changes in the dominant alleles result in decreases in the frequency of HCN, making populations especially susceptible to loss of HCN via drift.

In this paper, we develop a series of spatially-explicit, agent-based simulations to address the following question: (1) How do drift, gene flow, selection, and colonization history (i.e. urban to rural, rural to urban, always colonized) interact in the formation and maintenance of spatial clines in HCN? Our results show that stochastic changes in allele frequency can have the appearance of deterministic changes in the phenotype and we argue that this result extends beyond HCN to many phenotypes that are controlled by multiple genes.

**Table 1:** Direction of change in the frequency of HCN as the frequency of dominant alleles at underlying loci (i.e. *CYP79D15* or Li) either increases or decreases. In three out of four cases, HCN frequencies decrease. This makes populations particularly susceptible to loss of HCN via drift.

|  |  |  |
| --- | --- | --- |
|  | ***Li-***⇑ | ***Li*-**⇓ |
| ***CYP79D15-***⇑ | ⇑ | ⇓ |
| ***CYP79D15-***⇓ | ⇓ | ⇓ |