**Methods**

To examine the formation of spatial clines in HCN, we created a series of spatially-explicit simulations in Python 2.7 (CITE?) to track the frequency of HCN within populations across space. We set up a one-dimensional, linear matrix with 40 cells, consistent with the number of populations sampled across cities by Thompson et al. (2016), where each cell (hereafter patch) represents a patch of suitable habitat that can support a population of *Trifolium repens*. These simulations allowed for fine scale, independent control of both stochastic and deterministic parameters important for varying and maintaining the frequency of *CYP79D15* and *Li*—and thus HCN—in patches distributed across the landscape (Table 2). The order of events in the simulations are as follows: (1) Local reproduction (i.e. population growth), (2) selection, (3) migration, (4) colonization (Figure 1C). We first explored two broad colonization scenarios, described below, which differ in how they manipulate the amount of genetic drift acting within populations.

*Case 1:* *Gradient in carrying capacity across the matrix*

In the first scenario, we imposed a gradient in the carrying capacity (*K*) of populations across the matrix, thereby placing an upper-limit on the population size (*N,* Figure 1A). Drift is expected to be greatest in populations with the smallest carrying capacity and this method has been used in other agent-based simulations exploring the effects of drift, gene flow, and selection on patterns of local adaptation (Alleaume-Benharira et al. 2006). We first simulated a scenario where *N* is assumed to be greatest in rural populations (*N* = 1000) and decline linearly with increasing urbanization (*N* = 10 at urban end, Figure 1A). This scenario represents a case where clover populations were initially similar but increased fragmentation associated with urbanization reduced urban population sizes and increased the strength of drift. All 40 populations were initialized—and remained— at carrying capacity; thus, population growth is irrelevant in this first case. These simulations were run for 500 non-overlapping generations.