**Online supplementary materials for: Urban spandrels: the roles of genetic drift, gene flow and selection in the formation of parallel clines.**

**Contents of online supplementary materials:**

* Supplementary text:
  + S1: Effects of initial allele frequency variation on cyanogenesis cline formation
    - *Drift scenario 1:* *Gradient in carrying capacity across the matrix*
    - *Drift scenario 2:* *Colonization and founder events*
  + S2: Effects of selection on linkage between *CYP79D15* and *Li*
  + S3: Drift-selection balance through serial founder events (drift scenario 2)
  + S4: Effects of generation chosen for analysis on the formation and strength of clines in HCN
  + S5: Comparison of simulated slopes to standardized slopes for cyanogenesis clines observed across urban-rural gradients
* Supplementary figures S1 – S5

**Effects of initial allele frequency variation on cyanogenesis cline formation**

*Drift scenario 1:* *Gradient in carrying capacity across the matrix*

The initial frequency of both dominant alleles influenced the formation and strength of phenotypic clines in HCN. The strongest clines (β = 0.006) occurred when the frequency of both dominant alleles (i.e. *CYP79D15* and *Li*) was 0.5 (Figure S1A). The weakest clines occurred when the frequency of one or both dominant alleles was low (i.e. 0.1; 0.0005 < β < 0.002) whereas clines of intermediate strength occurred when either or both alleles were at high frequency (i.e. 0.9; 0.003 < β **<** 0.005; Figure S1A). These results hold regardless of the level of gene flow; increasing gene flow reduced the strength of clines, regardless of initial allele frequencies (Figure S1A).

The proportion of significantly positive clines was always greater than the proportion of negative clines, independent of initial allele frequencies. The greatest frequency of significantly positive clines peaked at 30% when the frequency of both dominant alleles was 0.5, followed by cases when one or both alleles were at low frequency (i.e. 0.1; 11 < % < 16) and finally by cases where one or both alleles were at high frequency (i.e. 0.9; 16 < % < 22; Figure S1B). Significantly negative clines were rare and only arose when the frequency of one or both dominant alleles was high (Figure S1B).

*Drift scenario 2:* *Colonization and founder events*

****

**Figure S1:** Effects of initial frequency of both dominant alleles (*CYP79D15* and *Li*) on (*a*) the mean strength of clines across 1000 simulations and (*b*) the proportion of significantly positive (open triangles with dashed line) and negative (black inverted triangles with solid line) clines. Simulations were run under a strong gradient in drift, manipulated by imposing a gradient in the maximum size of populations: rural populations were large (*N =* 1000) while urban populations were small (*N =* 10). In (*a*) we examined the mean slope of clines under no (open circles with dotted line), low (grey square, with dashed line), and high (black diamonds with solid line) gene flow. In (*b*) positive clines reflect significantly (*P <* 0.05) less HCN in urban populations relative to rural populations while negative clines reflect the opposite. All points represent means or proportions ± 95% confidence intervals.

**Effects of selection on linkage between *CYP79D15* and *Li***

We performed a small-scale simulation to examine the build-up of linkage disequilibrium (LD) between *CYP79D15* and *Li* when selection is acting on HCN. We initialized a single population with the frequency of both dominant alleles set to 0.5. From these allele frequencies, we calculated the frequency of all 16 possible diploid genotypes, assuming Hardy-Weinberg equilibrium. We then subjected these genotypes to selection, which acted against (negative selection coefficients) or in favor of (positive selection coefficients) cyanogenic genotypes. From the selected genotypes, we calculated the frequency of gametes, where heterozygotes were assumed to produce equal frequencies (i.e. 0.25) of all 4 possible gametes given the absence of physical linkage between *CYP79D15* and *Li* (i.e. recombination = 0.5). We calculated linkage disequilibrium as: r2 = D2 / (p*CYP79D15* × q*CYP79D15* × pLi ×qLi), where D represents the coefficient of linkage disequilibrium and is a function of gamete frequencies in any one generation. In the denominator, p*CYP79D15*, q*CYP79D15*, pLi, and qLi represent the frequency of dominant (i.e. p) and recessive (i.e. q) alleles at *CYP79D15* and *Li*. Thus, r2 is a measure of LD that accounts for allele frequencies and has a value of 1 when loci are in complete LD and 0 when they are in linkage equilibrium (i.e. independent of one another, REF). Gamete frequencies in following generations were then calculated from selected genotypes and this process was repeated recursively for 500 generations, allowing us to track the build-up of LD due to selection for cyanogenic and acyanogenic genotypes.

Our results show that the build-up of LD is minimal and decays rapidly over 500 generations (figure S2). Even under strong selection (e.g. -0.1), r2 reaches a maximum just under 0.0005, which is sufficiently close to zero to consider the loci in linkage equilibrium. Given these results, we ignored the build-up of LD due to selection in our simulations.



**Figure S2:** The build-up of linkage disequilibrium (LD) between *CYP79D15* and *Li* due to selection acting on cyanogenic white clover genotypes. Negative selection coefficients represent selection acting against cyanogenic clover genotypes while positive coefficients represent selection favouring cyanogenesis. Selection causes minimal build-up of LD between *CYP79D15* and *Li*, which decays rapidly over 500 generations.

**Drift-selection balance through serial founder events (drift scenario 2)**

Serial founder events from urban to rural populations constrained the ability of selection to generate strongly positive cyanogenesis clines. In the absence of founder events, increasing selection led to consistently stronger clines, independent of the rate of gene flow (figure S3*a*). When gene flow was low or absent, intermediate founder events (founding proportion = 0.2) resulted in the mean strength of clines being negative for all *s* < 0.025 (figure S3*b*), whereas the strongest positive clines occurred when *s* = 0.05 (β ≈ 0.005 for both low and no migration). High gene flow reduced the extent at which negative clines were formed by selection (figure 3*b*) and weaker selection (*s >* 0.005) was required before positive clines evolved in the presence of high gene flow. When founder effects were strong, selection had to be greater than 0.0025, 0.01, and 0.005 to generate positive clines when gene flow was absent (m = ), low and high, respectively (figure S3*c*). The strongest positive clines occurred when *s* = 0.05 (β ≈ 0.002 for no and low gene flow). These results are consistent with intermediate founder effects generating the strongest clines in HCN and further demonstrate that strong selection is required to overcome the formation of clines in the presence of an opposing drift gradient.

Serial founder events also influenced the extent to which selection generated positive and negative cyanogenesis clines. In the absence of founder events, and when selection is less than 0.005, both positive and negative clines occur with approximately 30% frequency (figure S3*d*). However, when *s* > 0.005, the frequency of positive clines rapidly increases to 100%, whereas negative clines declines to 0% (figure S3*d*). In the presence of intermediate founder events, negative clines are consistently more frequent for all *s <* 0.1, consistent with founder events of intermediate strength preferentially generating clines in HCN. However, when *s* ≥ 0.1, both positive and negative clines occur at less than 10% frequency (figure S3*e*). Finally, strong founder events result in little change in the frequency of positive clines, which fluctuate around 20% for all but the strongest selection coefficient (*s* = ) (figure S3*f*). By contrast, the frequency of negative clines rapidly decreases from 45% in the absence of selection to 0% when *s* ≥ 0.0025, becoming less common the negative clines when *s* ≥ 0.005 (figure S3*f*). These results further demonstrate that strong selection is required to overcome the formation of clines in HCN in the presence of an opposing gradients in genetic drift.



**Figure S3:** Serial founder events and selection interact in the formation of spatial clines in HCN. Populations colonized from the urban-most population to the rural-most population. Selection favours HCN+ genotypes in rural populations and HCN– genotypes in urban populations. (*a – c*) The mean slope of clines in HCN across 1000 simulations under no gene flow (open circles with dotted line), low gene flow (grey square with dashed line), or high gene flow (black diamonds with solid line). (*d* – f) The proportion of significantly positive (open triangles) and negative (black inverted triangle) clines across 1000 simulations. Founder events represent the proportion of alleles sampled in founding the new population and are either absent (proportion = 1.0, *a* and *d*), of intermediate strength (proportion = 0.2, *b* and *e*) or strong (proportion = 0.01, *c* and *f*). All points represent means or proportion ± 95% confidence intervals.

**Effects of generation chosen for analysis on the formation and strength of clines in HCN**

The generation chosen for analyses had little effect on our ability to assess the contributions of drift, gene flow, or selection on the formation and strength of phenotypic clines in HCN. For simplicity, we demonstrate this only for a strong gradient in carrying capacity (drift scenario 1, maximum rural *N* = 1000, minimum urban *N* = 10) under varying levels of gene flow. Following the formation of a cline, differences in the mean strength of clines across varying levels of gene flow remain consistent, independent of generation (figure S4*a*). The only exception to this is the strength of clines in the absence of gene flow, which decreases gradually over time due to drift, resulting in weaker clines at generation 500 than is evident at generation 250 (figure S4*a*). Nonetheless, this has no effects on our interpretation that increasing the amount of gene flow reduces the mean strength of clines (see main text figure 4*b*). Similarly, differences in the proportion of significantly positive clines remain qualitatively similar across generations (figure S4*b*). Therefore, the generation chosen for analysis has no influence on our ability to interpret the role of migration in influencing the formation and strength of cyanogenesis clines formed via drift.

****

**Figure S4:** Differences in the mean strength of clines (*a*) and the proportion of significantly positive clines (*b*) remain qualitatively similar regardless of which generation is chosen for analysis. (*a*) The mean strength of clines across 1000 simulations every generation from 1 to 500 under varying levels of gene flow. (*b*) The proportion of significantly positive clines across 1000 simulations every generation from 1 to 500 under varying levels of gene flow. Only levels of gene flow from 0 to 0.05 are shown to increase visibility of lines. Red bar at generation 250 represents the generation used in analyses and results are shown in main text (figures 4*b* and 4*d*).

**Comparison of simulated slopes to standardized slopes for cyanogenesis clines observed across urban-rural gradients**

We were interested in comparing the strength of clines produced by drift in our simulations to the strength of clines observed across urban-rural gradients in natural populations. For simplicity, we only examined the strength of clines simulated under weak or intermediate gradients in carrying capacity (drift scenario 1) and for no and high gene flow. Data for observed clines were obtained from [1] and [2]. Because the length of transects in our simulations and across cities varied, it was necessary first to standardize slopes before comparison. We standardized transects to a minimum value of 0 (urban-most population) and a maximum value of 1 (rural-most population). We then performed a linear regression using within-population HCN frequency as the response variable and standardized distance value as the predictor variable. Positive slopes represent less HCN in urban populations whereas negative slopes represent the opposite.

We found that the strength of observed clines is consistent with the strength of clines generated by drift in our simulations. The strongest simulated clines occurred under a strong gradient in drift in the absence of gene flow (–0.35 < βsimulated < 0.81, figure S5*a*). Increasing the amount of gene flow or decreasing the strength of the gradient in drift reduced the maximum strength of clines (figure S5*b* through S5*d*). The weakest simulated clines in HCN occurred under an intermediate gradient in drift with high gene flow (–0.24 < βsimulated < 0.27, figure S5*d*). The strength of observed clines ranged from –0.08 to 0.3 (figure S5*e*); thus, observed clines are within the range of even the weakest clines simulated under a gradient in drift, suggesting that drift is sufficient to generate clines as strong as those observed across replicated urbanization gradients.



**Figure S5:** Distribution of standardized slopes for simulated (*a* through d) and observed (e) cyanogenesis clines. Simulated slopes were generated using drift scenario 1 (i.e. gradient in carrying capacity) under varying les of gene flow and strengths of drift. (*a*) Slopes from simulations under a strong gradient in drift (minimum urban *N* = 10) and no gene flow. (*b*) Slopes from simulations under an intermediate gradient in drift (minimum urban *N* = 100) and no gene flow. (*c*) Slopes from simulations under a strong gradient in drift (minimum urban *N* = 10) and high gene flow (*m =* 0.05). (*d*) Slopes from simulations under an intermediate gradient in drift (minimum urban *N* = 100) and high gene flow (*m =* 0.05). (*e*) Distribution of slopes from urban-rural cyanogenesis clines observed across cities (n = 26) by Thompson et al. (2016) and Johnson et al. (2018). Blue bars represent clines that are significant at *P* < 0.05. Black dashed bar over histograms represents a slope of zero whereas the red dashed bar represents the mean slope.