Notes and Comments

Space-Time Relatedness and Hamilton's Rule for Long-Lasting Behaviors in Viscous Populations

Laurent Lehmann*

Morrison Institute for Population and Resource Studies, Stanford University, Stanford, California 94305

Submitted April 7, 2009; Accepted September 9, 2009; Electronically published November 13, 2009
Online enhancement: appendix.

ABSTRACT: Genes affect not only the behavior and fitness of their carriers but also that of other individuals. According to Hamilton's rule, whether a mutant gene will spread in the gene pool depends on the effects of its carrier on the fitness of all individuals in the population, each weighted by its relatedness to the carrier. However, social behaviors may affect not only recipients living in the generation of the actor but also individuals living in subsequent generations. In this note, I evaluate space-time relatedness coefficients for localized dispersal. These relatedness coefficients weight the selection pressures on long-lasting behaviors, which stem from a multigenerational gap between phenotypic expression by actors and the resulting environmental feedback on the fitness of recipients. Explicit values of space-time relatedness coefficients reveal that they can be surprisingly large for typical dispersal rates, even for hundreds of generations in the future.

Keywords: relatedness, Hamilton's rule, extended phenotype, niche construction.

> In the terms of the extended phenotype, alleles for larger lakes replaced alleles for smaller lakes. In the same terms, beavers can be said to carry within themselves genes whose phenotypic expression extends many miles away from the genes themselves. Why not hundreds of miles, thousands of miles? (R. Dawkins 1982)

Introduction

In evolutionary biology, a social behavior is a phenotype expressed by one individual that affects the fitness of another individual in a population (Hamilton 1964). For instance, parents transfer resources to their offspring, which help their juveniles to grow and survive. Workers in social insects raise the offspring of the queen. But conspecifics also compete for limited resources, and the gain in resources increasing the fitness of one individual is balanced by a loss of resources decreasing the fitness of others.

The selective pressure on a gene affecting a social behavior can conveniently be analyzed by using the concepts of inclusive fitness and evolutionary game theory (Hamilton 1964; Maynard-Smith 1982; Eshel 1996; Taylor 1996; Frank 1998; Rousset 2004; Grafen 2006). This usually consists of focusing on a mutant allele coding for a phenotypic feature whose value deviates by a small amount from that expressed by an individual bearing a resident (wild-type) allele and asking whether the mutant allele will spread in the gene pool. Whether the mutant allele will be favored by selection can then be ascertained by evaluating its inclusive fitness effect, which is the sum of the effects of the behavior of an individual bearing the mutant allele on the fitness of all recipients in the population, each weighted by the relatedness between the actor and the recipient (Hamilton 1964, 1970).

Relatedness measures the statistical association between the actor and the recipient breeding values (Michod and Hamilton 1980; Queller 1992; Frank 1997; Gardner et al. 2007). This often reduces to a measure of the extent to which the actor and recipient are more likely to share alleles identical by descent than are two individuals sampled at random from the population (Frank 1998). Relatedness can also be thought of as a ratio of two standardized transmission coefficients. It measures the extent to which the recipient is more likely to transmit the mutant allele to an offspring than is an individual sampled at random from the population, relative to the extent to which the actor is more likely to transmit the allele to an offspring than is a random individual. Relatedness thus describes a focal allele's valuation of another individual (in terms of transmission rate of replicate copies) relative to the individual in which it resides (Frank 1998, p. 58).

What is a focal allele's valuation of its transmission by an individual living in the distant future when its carrier reproduces and dies in the present? This question arises

* E-mail: laurent.lehmann@unine.ch.

Am. Nat. 2010. Vol. 175, pp. 136–143.
DOI: 10.1086/648554

from the observation that the effects of an organism on its environment are not necessarily completely and perfectly erased from one generation to the next (Darwin 1883; Lewontin 2000). Thus, phenotypic effects not only may change the fitness of recipients living in the generation of the actor but also may be felt by individuals living in the next or subsequent generations.

Phenotypic carryover effects across generations may be seen as a part of an individual's extended phenotype (Dawkins 1978, 1982), where genes extend their influence, beyond the bodily manifestations of the organism in which they reside and into the environment. Carryover effects can also be seen as niche-constructing phenotypes (Odling-Smee et al. 1996, 2003), which change the structure and configuration of the environment of future generations that may then feed back on the evolution of the phenotypes. In the presence of a multigenerational gap between behavioral modification of the environment by an actor and the fitness consequences on recipients, there must be a covariance between the actor and the recipients' phenotypes across generations in order for long-lasting traits to be affected by natural selection (Brodie 2005).

The aim of this note is to introduce and evaluate intergenerational relatedness coefficients for finite populations following an isolation by distance pattern of population structure. Such relatedness coefficients provide a measure of phenotypic covariances across generations. They quantify the extent to which selection pressures in downstream generations affect the evolution of long-lasting behaviors and how this varies with variation in demographic assumptions. These relatedness coefficients can then be used in Hamilton's rule in order to assess the direction of selection on long-lasting behaviors.

This work is an addendum to an inclusive fitness analysis of the evolution of long-lasting phenotypes (Lehmann 2007, 2008; Sozou 2009), which I complement here by exploring specifically and explicitly the stationary distribution of intergenerational relatedness under localized dispersal. The results suggest that such relatedness coefficients can take substantial values for hundreds of generations for dispersal rates typical of natural populations.

Model

Biological Outline

Assume a haploid population with n_d demes, each with a constant number N of adult individuals. The structure of this population is assumed to follow the standard homogeneous isolation by distance model (e.g., Maruyama 1970; Malécot 1975; Nagylaki 1983; Rousset 2004). That is, demes can be thought of as points in a homogeneous discrete space (e.g., equally spaced points on a circle if the habitat is one dimensional) and are connected by dispersal. The dispersal distribution of individuals is assumed to be symmetric (in a one-dimensional habitat, individuals move clockwise and counterclockwise with the same distributions) and identical for all demes, but it can otherwise take any form (geometric, binomial, etc.). Finally, the exogenous environment of the population (all biotic and abiotic factors exogenous to the population) is assumed to be constant and identical at all times.

Consider that a given phenotype expressed by an individual in this population may affect the survival or the fecundity (vital rates) of any other individual living in the present or in the future of the population. Examples of long-lasting phenotypic effects may include the construction of a nest or a burrow, the consumption of a resource, or the emission of detritus and may thus result in an endogenous change of the environment. Long-lasting phenotypic effects are assumed here to follow two crucial homogeneity features.

Spatial homogeneity of phenotypic effects. This means that the distribution of the phenotypic effects of a behavior expressed by a focal individual living at any focal lattice point (focal deme) on the fitness of others living at the same or other lattice points is identical at all lattice points (for more details, see Rousset and Billiard 2000; Lehmann 2008). With this assumption and that of the homogeneity of population structure, an integer coordinate k (single in one dimension, a pair in two dimensions, etc.) can be used to measure the "spatial distance" of a deme relative to the position of the focal deme. In a one-dimensional habitat, $k = 0, 1, 2, \ldots$ can be thought of as the number of steps on the lattice separating two demes.

Temporal homogeneity of phenotypic effects. This means that the distribution of the phenotypic effects of a behavior expressed by a focal individual at any focal time point (focal generation) on the fitness of other individuals living at future time points is identical at all times. With this, an integer coordinate $t = 0, 1, 2, \ldots$ can be used to measure the "temporal distance" of a deme relative to the focal deme ($t = 0$ stems for the focal generation; for more details, see Lehmann 2008).

Space-Time Relatedness and Hamilton's Rule

When gene action is additive and selection is weak in the population described above, the condition under which a mutant allele with long-lasting phenotypic effects is favored by selection over a wild-type allele can be expressed in terms of Hamilton's rule as

$$\sum_t \sum_k R_{k,t} b_{k,t} - c > 0 \qquad (1)$$

(eqq. [A1]–[A6] in the appendix in the online edition of the *American Naturalist*). This invasion condition depends on three quantities. The first quantity is the net change $-c$ in the fitness of a focal individual living in a focal generation and stemming from the focal individual expressing the mutant allele during its whole life span. The second quantity is the change $b_{k,t}$ in the fitness of the whole set of individuals living at distance k from the focal deme at t generations after the focal generation (referred to as deme k, t) and stemming from the focal individual expressing the mutant allele. The third quantity is the relatedness coefficient

$$R_{k,t} = \lim_{\mu \to 0} \frac{Q_{k,t} - \overline{Q}_t}{1 - \overline{Q}_0} \tag{2}$$

between the focal individual and a recipient from deme k, t, where μ is the mutation rate of the gene underlying the behavior (assumed to follow the infinite allele model; Kimura and Crow 1964); $Q_{k,t}$ is the stationary probability that an allele sampled in the actor is identical by descent with a homologous allele sampled in an individual chosen at random from deme k, t; and $\overline{Q}_t = \sum_k Q_{k,t}/n_d$ is the average probability of identity between two homologous alleles sampled in two distinct individuals living at t generations apart (for diploid organism, the 1 in the denominator of eq. [2] has to be replaced by the coancestry with self).

The relatedness coefficient $R_{k,t}$ (eq. [2]) provides a measure of the extent to which an individual sampled in deme k, t is more (or less) likely to transmit a mutant allele to the next generation than is a randomly sampled individual from t, relative to the extent to which the focal individual is more likely to transmit the mutant allele to the next generation than is another individual sampled in its generation. The definition of $R_{k,t}$ also entails that the average relatedness of a focal individual with individuals living at t generations from the focal generation is equal to 0: $\overline{R}_t = \sum_k R_{k,t}/n_d = 0$. Hence, for each t, the total number of individuals that are positively related to the focal individual, each weighted by their relatedness to the focal individual, is exactly equal to the total relatedness weighted sum of individuals that are negatively related to him.

The relatedness coefficient $R_{k,t}$ is akin to the classical measure of population structure, F_{ST} (Wright 1951; Crow and Aoki 1984; Cockerham and Weir 1987; Slatkin 1991), with the only difference that it takes the temporal differentiation between pairs of demes into account in addition to the classic spatial differentiation. But it is important to note that $R_{k,t}$ is a stationary value of relatedness, where t indexes the number of generations between the sampling of two individuals in the population. Hence, it is not the same as the transient value of relatedness (or n_d) between two individuals sampled in the same generation at distance k. As such, $R_{k,t}$ is different from the time dynamics of the within-generation probabilities of identity by descent that were analyzed in previous work (e.g., Malécot 1975; Varvio et al. 1986).

Because the relatedness coefficient $R_{k,t}$ weights the fitness change $b_{k,t}$ of individuals living in deme k, t, it quantifies the importance of the selection coefficient $b_{k,t}$ for the evolution of the mutant allele. Equation (1) gives the generic invasion condition for a mutant allele in a homogeneous population of constant size without class structure (e.g., age or stage structure). As such, no competitive effect has been scaled out in equation (1), and the values both $b_{k,t}$ and $R_{k,t}$ can take depend on additional details of the life history. In "Space-Time Relatedness for Wright-Fisher Reproduction," I provide examples of values that $R_{k,t}$ can take under particular demographic assumptions and return to the issue of the $b_{k,t}$ coefficients in "Discussion."

Space-Time Relatedness for Wright-Fisher Reproduction

The relatedness $R_{k,t}$ is analyzed here under a Wright-Fisher reproductive scheme (Ewens 2004), which corresponds to a semelparous life cycle and which underlies classical isolation by distance models (e.g., Maruyama 1970; Malécot 1973, 1975; Nagylaki 1983; Epperson 1999; Rousset 2004). For ease of presentation, I consider only a one-dimensional habitat, where the events of the life cycle are as follows. (1) Each of the N adult individuals in a deme produces a very large number of juveniles. After reproduction, all adults die. (2) Each juvenile either remains philopatric with probability m_0 or may disperse with probability m_k to a deme at $k = 1, 2, \ldots, n_d - 1$ steps away from its natal deme, either clockwise (probability 1/2) or counterclockwise (probability 1/2).

For this life cycle, the relatedness between a focal individual and a recipient living at $k = 0, 1, 2, \ldots$ steps away from the focal deme at $t = 1, 2, 3, \ldots$ generations after the focal generation can be expressed as

$$R_{k,t} = \frac{1}{Nn_d + G} \sum_{h=1}^{n_d - 1} \frac{\psi_h^t}{1 - \psi_h^2} e^{-ikz(h)}, \tag{3}$$

where $i = (-1)^{1/2}$, $z(h) = 2\pi h/n_d$, $G = \sum_{h=1}^{n_d - 1} \psi_h^2/(1 - \psi_h^2)$, and the scalars ψ_h are the eigenvalues of the migration matrix (eqq. [A7]–[A18]). The relatedness coefficient between two individuals sampled at k steps apart in the same generation ($t = 0$) is given by substituting $t = 2$ into equation (3); that is, $R_{k,0} = R_{k,2}$ (for details, see text below; eqq. [A9], [A18]). Assuming that there is a positive level of philopatry ($m_0 > 0$) and that a line of descent of a gene residing in a focal deme is eventually able to reach

every deme in the population in the long run, the migration matrix has a single unit eigenvalue, $\psi_0 = 1$, and all eigenvalues appearing in equation (2) are strictly <1 in absolute value (see paragraph below; eq. [A9]).

It follows from these considerations that the stationary space-time relatedness decays with time and its long-term value is equal to 0: $\lim_{t\to\infty} R_{k,t} = 0$ for all k. This result is expected to hold more generally in finite populations, regardless of the details of the life cycle, because when looking sufficiently far in the future (or the past), the line of descent (or the ancestral line) of a gene will be distributed uniformly in the whole population (instead of being clustered spatially and temporally around the focal deme). But how fast does the space-time population structure decay for a given dispersal distribution?

Explicit values of $R_{k,t}$ will be computed by using a geometric dispersal distribution (truncated because there are a finite number n_d of demes) defined as

$$m_i = \begin{cases} 1-m & \text{when } i = 0 \\ m\dfrac{(1-g)g^{i-1}}{1-g^{(n_d-1)}} & \text{for } 0 < i \le n_d - 1 \text{ and } 0 \text{ otherwise} \end{cases}, \tag{4}$$

where m is the migration probability out of the natal deme. With this, an individual remains in its natal deme with probability $1-m$ or disperses with probability m in a geometric manner up to $n_d - 1$ steps apart, either clockwise or counterclockwise on the lattice. Figure 1 illustrates that the parameter g (varying between 0 and 1) in equation (4) allows one to investigate a continuum of spatial structures, ranging from the stepping-stone model of dispersal when $g \to 0$ (in which case dispersal is localized, $m_1 = m$ and $m_i = 0$ for $i > 1$; Kimura and Weiss 1964) to Wright's (1931) island model of dispersal when $g \to 1$ (in which case dispersal is random, with probability $m/(n_d - 1)$ for any nonnatal deme).

The relatedness coefficients $R_{k,t}$ are shown in figure 2 as a function of the spatial distance between the two demes in which individuals are sampled for various values of the temporal distance between them. This figure illustrates two typical patterns of relatedness between pairs of individuals sampled from the same generation under isolation by distance. First, relatedness between individuals decreases as the spatial distance between them increases (e.g., Kimura and Weiss 1964; Malécot 1975; Nagylaki 1983; Hartl and Clark 2007). Second, relatedness takes positive values for individuals sampled in adjacent demes and negative values for individuals sampled in demes far apart from each other (Grafen 2007). Hence, both increasing the vital rates of individuals living nearby and decreasing those of individuals living far away may increase inclusive fitness. Figure 2 also illustrates that these two features hold more generally for all temporal distances and with the isolation by distance pattern decreasing as the temporal distance increases, which follows from the temporal decay of the stationary space-time probabilities of identity by descent (the $Q_{k,t}$'s; Malécot 1973; Epperson 1999).

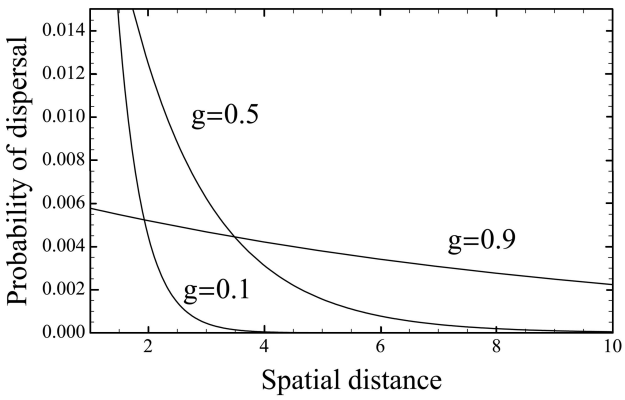


Figure 1: Geometric dispersal distribution given by equation (4) for $m = 0.1$ and for various values of the parameter g, which parametrizes the range of dispersal. One has $g = 0.1$ for the most rapidly decreasing curve in both panels (dispersal is mainly local), $g = 0.9$ for the least rapidly decreasing curve in both panels, and $g = 0.5$ for the intermediate curve. The figure illustrates the dispersal distribution when an individual moves clockwise on the lattice, with its natal deme located at the origin on the horizontal axis with $n_d = 20$.

However, figures 2–4 reveal a biologically more interesting point: namely, that the relatedness between individuals sampled in different generations from the same or adjacent demes remains high for up to hundreds of generations for migration rates typical of natural populations (for all figures, two migrants per generation were assumed; i.e., $Nm = 2$; Barton 2001, p. 334; Hartl and Clark 2007, their table 6.5, p. 302). Only when the dispersal rate becomes high and the dispersal distribution tends to that of the island model with random migration ($g \to 1$ in eq. [4]) does the temporal clustering between individuals sampled in neighboring demes vanish.

The high relatedness between a pair of individuals sampled at many generations apart in the same or in adjacent demes (figs. 2–4) is explained by the fact that under localized dispersal individuals tend to move to nearby demes (e.g., stepping stone dispersal), which results in a high probability that the pair of individuals sampled involves both an ancestor and its descendant. This probability will be much lower under random migration because once a line of descent of an ancestor moves out of its deme, it has a much lower probability to return to the deme of the ancestor (or an adjacent deme) in later generations than under localized dispersal. Hence, localized dispersal may lead to high spatiotemporal clustering of gene lineages.

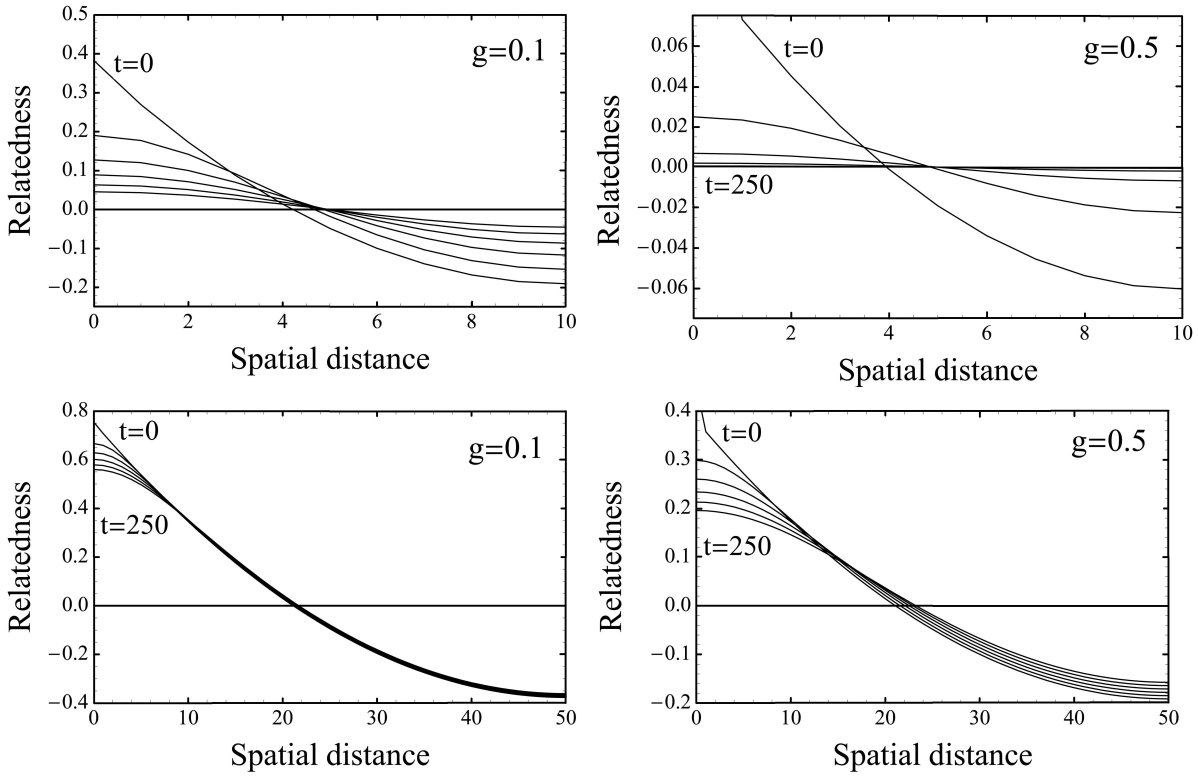


Figure 2: Space-time relatedness coefficient $R_{k,t}$ (eq. [2]) graphed as a function of the spatial distance k between two individuals for various values of the temporal distance t between them for the geometric dispersal distribution (eq. [4]) with $N = 20$ and $m = 0.1$. From the highest to the lowest curve on the left-hand side of the graphs, $t = 0$, 50, 100, 150, 200, and 250. The top row is for $n_d = 20$, while the bottom row is for $n_d = 100$; in both cases, the left-hand graph has $g = 0.1$, and the right-hand graph has $g = 0.5$. The figure illustrates that two individuals sampled at hundreds of generations apart can be strongly related when they are sampled in adjacent demes and when dispersal is localized. When dispersal becomes less localized (shifting from $g = 0.1$ to $g = 0.5$), both the spatial and temporal population structures decrease.

Discussion

The model presented in this note is based on some important simplifying assumptions, in particular, weak selection, additive gene action, constant exogenous environment, and constant population size. The model is thus not meant to be a realistic description of extended phenotypes or niche construction. A more realistic model would involve dynamical demographic variables (such as deme sizes and conditions) varying in both space and time jointly with phenotypes generated by the expression of interacting loci within individuals across multitudes of generations. Such features are likely to disturb the dynamics of space-time relatedness from its neutral trajectory on which the current results rely. Nevertheless, the simplifying assumptions used here allow one to contrast directly short-term and long-term phenotypic effects of behaviors. This is because space-time relatedness coefficients capture the relevance of selection pressures in downstream generations for the evolution of a focal phenotype expressed in the present (i.e., how much $b_{k,t}$ for $t > 0$ will affect selection on the mutant allele; see eq. [1]).

As is usually the case for relatedness coefficients (or, more generally, F statistics; e.g., Wright 1951; Crow and Aoki 1984; Cockerham and Weir 1987; Slatkin 1991), the space-time relatedness coefficients, $R_{k,t}$, involve the identity between pairs of genes sampled in different classes of individuals (eq. [2]) and can thus be estimated from genetic markers (e.g., Queller and Goodnight 1989; Chapuisat et al. 1997; Fontanillas et al. 2004). However, because $R_{k,t}$ is defined relative to a population average probability of identity (i.e., $\overline{Q}_t$ in eq. [2]), estimation of average gene frequency in the population is required in order to evaluate it, which in practice may be complicated or infeasible. In order to circumvent this issue, one could define space-time relatedness coefficients relative to the probability of identity between pairs of genes sampled in the focal deme in the focal generation (and thus replace both $\overline{Q}_t$ and $\overline{Q}_0$ with $Q_{0,0}$ in eq. [2], which can still be

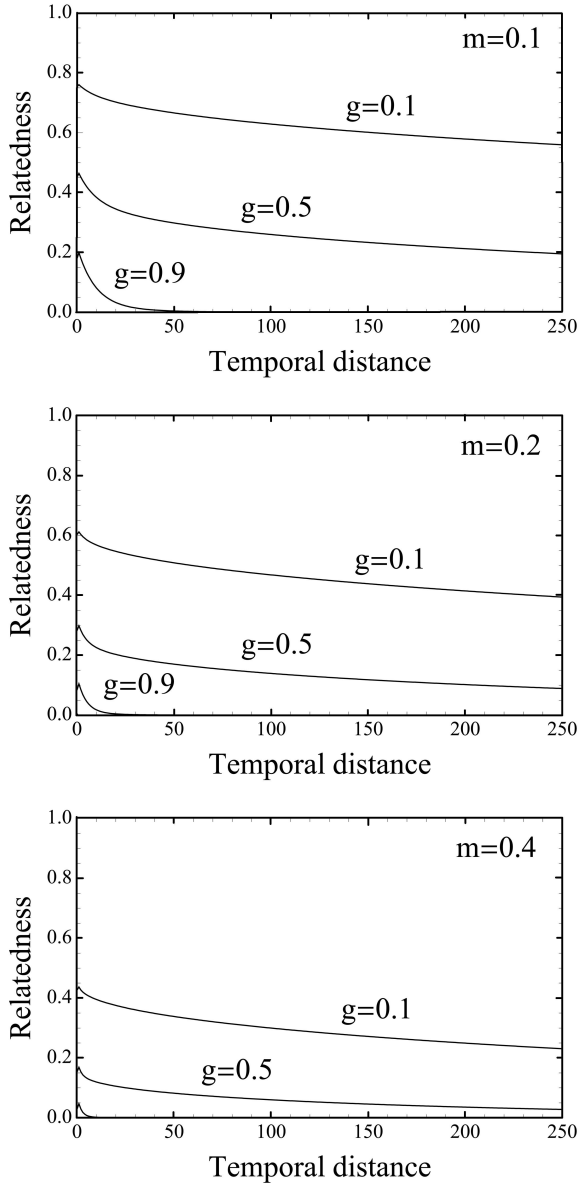


Figure 3: Space-time relatedness coefficient $R_{0,t}$ between two individuals sampled from the same deme as a function of the number of generations t between them for $N = 20$, $n_d = 100$, and various combinations of g and m. In each graph, the top curve is for $g = 0.1$ (dispersal is mainly local), the middle curve is for $g = 0.5$, and the bottom curve is for $g = 0.9$, while from the top to bottom graph, one has $m = 0.1$, 0.2, and 0.4. Hence, when dispersal becomes less localized (g increases), the temporal structure vanishes more rapidly, and an increase in migration, m, has exactly the same effect.

used in eq. [1] in order to assess the direction of selection because of the property that fitness effects sum to 0; eq. [A3]). This alternative definition of relatedness is maybe less intuitive for understanding the selective pressure on long-lasting behaviors (because relatedness to patch mates is 0), but it has an advantage from an empirical point of view since it allows one to evaluate relatedness using local allele frequencies only (Rousset 2002, 2006).

Finally, it is worth recalling that while the condition under which a mutant allele is favored by selection depends on both the relatedness coefficients ($R_{k,t}$) and the selection coefficients ($-c$ and $b_{k,t}$ in eq. [1]), the latter have not been evaluated here explicitly (specific examples of $b_{k,t}$ coefficients are presented in Lehmann 2008 and Sozou 2009). Selection coefficients stemming from the expression of social behaviors in spatially subdivided populations depend considerably on the details of the biological situation under focus, such as the timing of social interactions and the mode of competition, among a plethora of other life cycle features (e.g., Taylor 1992*a*; West et al. 2002, 2007; Rousset 2004; Lion and van Baalen 2007; Grafen and Archetti 2008).

Even for the life cycle assumptions leading to the specific examples of relatedness given by equation (3), there are a wide variety of possible outcomes for the values the selection coefficients $b_{k,t}$ can take. There are situations where the present and the long-lasting indirect effects ($b_{k,0}$ and $b_{k,t}$ for $t > 0$, respectively) will be null, such as when density-dependent competition occurs completely before dispersal (Wade 1985). There is also the classic situation were the present indirect effects, $b_{k,0}$, will take values such as to cancel the present generation benefits to kin (Taylor 1992*a*, 1992*b*) but in which case the presence of long-lasting indirect effects, $b_{k,t}$ for $t > 0$, allow for selection on altruism due to future generations benefits to kin (Lehmann 2008). Finally, there are scenarios where both present and future indirect effects may be large and may markedly affect the selective pressure on the mutant allele, such

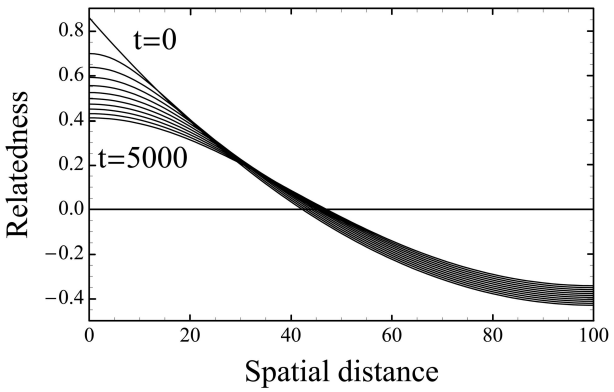


Figure 4: Space-time relatedness coefficient $R_{k,t}$ graphed as a function of k for various values of t, ranging from $t = 0$ to $t = 5{,}000$ in steps of 200 generations for $m = 0.1$, $g = 0.1$, $N = 20$, and $n_d = 200$.

as when the behavior primarily increases the number of dispersing individuals (Rogers 1990), in which case the effect of the behavior on present and future local competition is negligible.

Regardless of the exact magnitude of the indirect selection coefficients, the main point to take away is that if they are nonnull ($b_{k,t} \neq 0$ for $t > 0$; i.e., there is ecological inheritance sensu Odling-Smee et al. 1996, 2003), then behaviors may be considerably shaped by delayed effects under localized dispersal in natural populations. Indeed, the values taken by the space-time relatedness coefficients suggest that it is plausible that long-lasting phenotypes are subject to selection even if there is a gap of many generations between the behavioral modification of the environment and the fitness consequences on recipients (figs. 2–4). Why not hundreds of generations, thousands of generations?

Acknowledgments

I am grateful to J. Granka and J. Van Cleeve for carefully reading the manuscript and correcting the English and to D. Weissman for his many comments that improved the presentation. I also thank F. Rousset for helpful discussions. This work was supported by a grant from the Swiss National Science Foundation to L.L. and by National Institutes of Health grant GM28016 to M. W. Feldman.

Literature Cited

Barton, N. H. 2001. The evolutionary consequences of gene flow and local adaptation: future approaches. Pages 329–340 *in* J. Clobert, E. Danchin, A. A. Dhondt, and J. D. Nichols, eds. Dispersal. Oxford University Press, Oxford.

Brodie, E. D. 2005. Caution: niche construction ahead. Evolution 51: 249–251.

Chapuisat, M., J. Goudet, and L. Keller. 1997. Microsatellites reveal high population viscosity and limited dispersal in the ant *Formica paralugubris*. Evolution 51:475–482.

Cockerham, C. C., and B. S. Weir. 1987. Correlations, descent measures: drift with migration and mutation. Proceedings of the National Academy of Sciences of the USA 84:8512–8514.

Crow, J. F., and K. Aoki. 1984. Group selection for a polygenic behavioral trait: estimating the degree of population subdivision. Proceedings of the National Academy of Sciences of the USA 81: 6073–6077.

Darwin, C. 1883. The formation of vegetable mold through the action of worms, with observations on their habits. J. Murray, London.

Dawkins, R. 1978. Replicator selection and the extended phenotype. Zeitschrift für Tierpsychologie 47:61–76.

———. 1982. The extended phenotype. Oxford University Press, Oxford.

Epperson, B. K. 1999. Gene genealogies in geographically structured populations. Genetics 152:797–806.

Eshel, I. 1996. On the changing concept of evolutionary population stability as a reflection of a changing point of view in the quantitative theory of evolution. Journal of Mathematical Biology 34: 485–510.

Ewens, W. J. 2004. Mathematical population genetics. Springer, New York.

Fontanillas, P., E. Petit, and N. Perrin. 2004. Estimating sex-specific dispersal rates with autosomal markers in hierarchically structured populations. Evolution 58:886–894.

Frank, S. A. 1997. The Price equation, Fisher's fundamental theorem, kin selection, and causal analysis. Evolution 51:1712–1729.

———. 1998. Foundations of social evolution. Princeton University Press, Princeton, NJ.

Gardner, A., S. A. West, and N. Barton. 2007. The relation between multilocus population genetics and social evolution theory. American Naturalist 169:207–226.

Grafen, A. 2006. Optimization of inclusive fitness. Journal of Theoretical Biology 238:541–563.

———. 2007. An inclusive fitness analysis of altruism on a cyclical network. Journal of Evolutionary Biology 20:2278–2283.

Grafen, A., and M. Archetti. 2008. Natural selection of altruism in inelastic viscous homogeneous populations. Journal of Theoretical Biology 252:694–710.

Hamilton, W. D. 1964. The genetical evolution of social behaviour. I. Journal of Theoretical Biology 7:1–16.

———. 1970. Selfish and spiteful behavior in an evolutionary model. Nature 228:1218–1220.

Hartl, D., and A. G. Clark. 2007. Principles of population genetics. 4th ed. Sinauer, Sunderland, MA.

Kimura, M., and J. F. Crow. 1964. The number of alleles that can be maintained in a finite population. Genetics 49:725–738.

Kimura, M., and G. H. Weiss. 1964. The stepping stone model of population structure and the decrease of genetic correlation with distance. Genetics 49:561–576.

Lehmann, L. 2007. The evolution of trans-generational altruism: kin selection meets niche construction. Journal of Evolutionary Biology 20:181–189.

———. 2008. The adaptive dynamics of niche constructing traits in spatially subdivided populations: evolving posthumous extended phenotypes. Evolution 62:549–566.

Lewontin, R. C. 2000. The triple helix: genes, organism, and environment. Harvard University Press, Cambridge, MA.

Lion, S., and M. van Baalen. 2007. Self-structuring in spatial evolutionary ecology. Ecology Letters 11:277–295.

Malécot, G. 1973. Génétique des population diplodes naturelle dans le cas d'un seul locus. III. Parenté, mutations et migration. Annale de Génétique et de Selection Animale 5:333–361.

———. 1975. Heterozygosity and relationship in regularly subdivided populations. Theoretical Population Biology 8:212–241.

Maruyama, T. 1970. Effective number of alleles in a subdivided population. Theoretical Population Biology 1:273–306.

Maynard-Smith, J. 1982. Evolution and the theory of games. Cambridge University Press, Cambridge.

Michod, R. E., and W. D. Hamilton. 1980. Coefficients of relatedness in sociobiology. Nature 288:694–697.

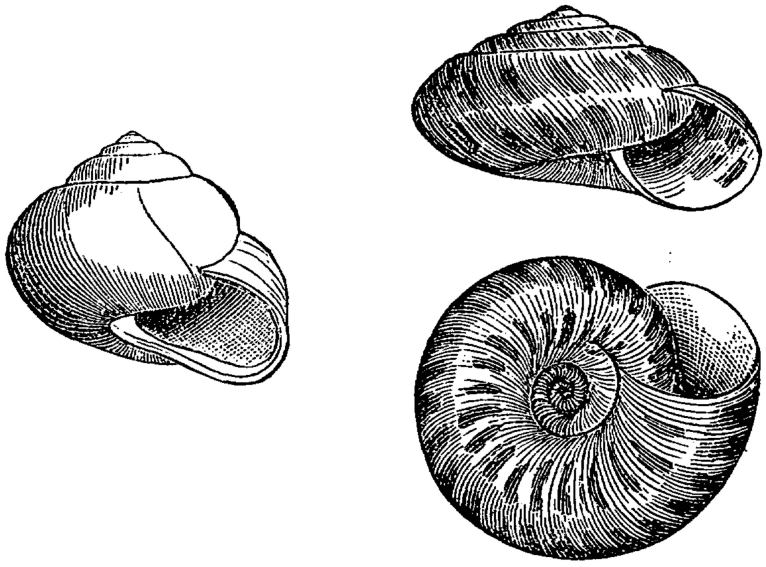
Nagylaki, T. 1983. The robustness of neutral models of geographical variation. Theoretical Population Biology 24:268–294.

———. 1998. The expected number of heterozygous sites in a subdivided population. Genetics 149:1599–1604.

Odling-Smee, F. J., K. L. Laland, and M. W. Feldman. 1996. Niche construction. American Naturalist 147:641–648.

———. 2003. Niche construction. Princeton University Press, Princeton, NJ.
Queller, D. C. 1992. A general model for kin selection. Evolution 46: 376–380.
Queller, D. C., and K. F. Goodnight. 1989. Estimating relatedness using genetic markers. Evolution 43:258–275.
Rogers, A. R. 1990. Group selection by selective emigration: the effects of migration and kin structure. American Naturalist 135:398–413.
Rousset, F. 2002. Inbreeding and relatedness coefficients: what do they measure? Heredity 88:371–380.
———. 2004. Genetic structure and selection in subdivided populations. Princeton University Press, Princeton, NJ.
———. 2006. Separation of time scales, fixation probabilities and convergence to evolutionarily stable states under isolation by distance. Theoretical Population Biology 69:165–179.
Rousset, F., and S. Billiard. 2000. A theoretical basis for measures of kin selection in subdivided populations: finite populations and localized dispersal. Journal of Evolutionary Biology 13:814–825.
Slatkin, M. 1991. Inbreeding coefficients and coalescence times times. Genetical Research 58:167–175.
Sozou, P. D. 2009. Individual and social discounting in a viscous population. Proceedings of the Royal Society B: Biological Sciences 276:2955–2962.
Taylor, P. D. 1992*a*. Altruism in viscous populations: an inclusive fitness model. Evolutionary Ecology 6:352–356.
———. 1992*b*. Inclusive fitness in a homogeneous environment. Proceedings of the Royal Society B: Biological Sciences 240:299–302.
———. 1996. Inclusive fitness arguments in genetic models of behaviour. Journal of Mathematical Biology 34:654–674.
Varvio, S. L., R. Chakraborty, and M. Nei. 1986. Genetic variation in subdivided populations and conservation genetics. Heredity 57: 189–198.
Wade, M. J. 1985. Soft selection, hard selection, kin selection, and group selection. American Naturalist 125:61–73.
West, S. A., I. Pen, and A. S. Griffin. 2002. Cooperation and competition between relatives. Science 296:72–75.
West, S. A., A. S. Griffin, and A. Gardner. 2007. Evolutionary explanations for cooperation. Current Biology 17:661–672.
Wright, S. 1931. Evolution in Mendelian populations. Genetics 16: 97–159.
———. 1951. The genetical structure of populations. Annals of Eugenics 15:323–354.

Associate Editor: Ross H. Crozier
Editor: Donald L. DeAngelis



Left, *Helix hortensis* Muller. "It is unquestionably identical with the European species and is supposed to have found its way to this country through commercial intercourse, though it seems strange that, while in the old country it is found near the habitations of men, in this country it occurs only upon the most uninhabitable islands." *Right*, *Helix alternate* Say. "On islands, they often occur in the greatest profusion. When in captivity, they lie buried most of the time under the moist earth, and appear to suffer more from the want of moisture than other species." From "The Land Snails of New England (Continued)" by Edward S. Morse (*American Naturalist*, 1867, 1:186–188).

 DOI: 10.1086/648554

Appendix from L. Lehmann, "Space-Time Relatedness and Hamilton's Rule for Long-Lasting Behaviors in Viscous Populations"

(Am. Nat., vol. 175, no. 1, p. 136)

Supplemental Information

Hamilton's Rule and Space-Time Relatedness

The expression for Hamilton's rule (eq. [1]) will be derived from previous inclusive fitness theory results for spatially subdivided populations.

Inclusive Fitness Effect

When gene action is additive and selection is weak in a homogeneous population, the change of the fixation probability (relative to neutrality) of a single mutant allele, whose phenotypic value is different from the phenotype resulting from the expression of a wild-type allele, can be written as

$$\phi = \lim_{\mu \to 0} \frac{S}{1 - Q_{0,0}}, \tag{A1}$$

where μ is the mutation rate of the gene underlying the behavior (taken here to follow the infinite allele model; Kimura and Crow 1964); $Q_{0,0}$ is the stationary probability of identity by descent between a pair of homologous genes sampled from two individuals from the same deme and from the same generation, which is evaluated in a neutral model (no selection); and S is the gradient of selection on the mutant allele, which can be expressed as an inclusive fitness effect (Rousset and Billiard 2000; Rousset 2004).

For the demographic setting introduced in the main text ("Biological Outline"), the gradient of selection can be expressed as

$$S = -c + \sum_{t} \sum_{k} b_{k,t} Q_{k,t}, \tag{A2}$$

where $-c$ is the change in the fitness of the focal individual stemming from the focal expressing the mutant allele, $b_{k,t}$ is the change in the fitness of the focal individual stemming from the expression of the mutant allele by the whole set of individuals living at distance k from the focal deme at t generations before the focal generation, and $Q_{k,t}$ is the stationary probability that a gene sampled in the focal individual is identical by descent with a homologous gene sampled in an individual chosen at random from deme k,t (Lehmann 2007, eq. [A11]; 2008, eq. [4]).

The homogeneity assumptions of the model entail that all individuals in the population face the same set of problems at all times (strategic equivalence; Grafen 2006, p. 543). The actor and the recipients can thus be interchanged, and one can use a future oriented interpretation of the inclusive fitness effect S. Then, $b_{k,t}$ can be interpreted as the effect of a focal individual on the fitness of the whole set of individuals living in deme k at t generations after the focal generation, and $Q_{k,t}$ gives the probability that a recipient carries a homologous gene identical by descent to that of the focal individual (for more details, see Lehmann 2008). If the gradient of selection is positive ($S > 0$), selection favors the mutant allele, which then has a higher probability of fixation than a neutral allele (or that of a wild-type allele in a population of mutants).

Inserting Relatedness

The condition of invasion of the mutant allele obtained from equation (A2) will now be expressed in terms of the relatedness coefficients $R_{k,t}$ defined in equation (2). To that end, we will use the result that the fitness changes appearing in equation (A2) sum to 0 in the following way:

$$-c + \sum_k b_{k,0} = 0,$$
$$\sum_k b_{k,t} = 0 \text{ for all } t \geq 1. \tag{A3}$$

These equalities follow from the fact that the changes in the fitness of a focal individual resulting from all individuals living in the population at t generations before the focal generation and all expressing the mutant allele must sum up to 0. Otherwise, a systematic change in gene frequency would occur in a monomorphic population, that is, under neutrality ($Q_{k,t} = 1$ for all k and all t), which is not possible (i.e., fitness effects are zero sum).

Inserting equation (A2) into equation (A1), we have

$$\phi = \lim_{\mu \to 0} \frac{1}{1 - Q_{0,0}} \left(-c + \sum_k b_{k,0} Q_{k,0} + \sum_{t \geq 1} \sum_k b_{k,t} Q_{k,t} \right), \tag{A4}$$

and by subtracting $\overline{Q}_0(-c + \sum_k b_{k,0})$ from the first two terms in parentheses and $\sum_{t \geq 1} \overline{Q}_t \sum_k b_{k,t}$ from the third term, one obtains

$$\phi = \lim_{\mu \to 0} \frac{1 - \overline{Q}_0}{1 - Q_{0,0}} \left[-c + \sum_k b_{k,0} \left(\frac{Q_{k,0} - \overline{Q}_0}{1 - \overline{Q}_0} \right) + \sum_{t \geq 1} \sum_k b_{k,t} \left(\frac{Q_{k,t} - \overline{Q}_t}{1 - \overline{Q}_0} \right) \right]. \tag{A5}$$

On substitution of the definition of $R_{k,t}$ (eq. [2]) into this equation, one has

$$\phi = \lim_{\mu \to 0} \frac{1 - \overline{Q}_0}{1 - Q_{0,0}} \left(\sum_t \sum_k R_{k,t} b_{k,t} - c \right), \tag{A6}$$

and the mutant allele is favored by selection when the term in parentheses is positive, which is equation (1).

Space-Time Relatedness

Here $R_{k,t}$ will be expressed in terms of the dispersal distribution for the life cycle described in the main text.

Probabilities of Identity by Descent

A classical result is that the stationary probability of identity $Q_{k,0}$ between two individuals sampled in the same generation at k steps apart on a one-dimensional circular lattice for the life cycle described in the main text can be written as

$$Q_{k,0} = \frac{1 - Q_{0,0}}{N n_{\mathrm{d}}} \sum_{h=0}^{n_{\mathrm{d}}-1} \frac{\gamma \lambda_h}{1 - \gamma \lambda_h} e^{-ikz(h)} \tag{A7}$$

(Malécot 1973, eq. [15]; 1975, eq. [18]; Epperson 1999, eq. [B4]; Rousset 2004, chap. 3), where $\gamma = (1 - \mu)^2$, $i = (-1)^{1/2}$, $z(h) = 2\pi h / n_{\mathrm{d}}$, and $\lambda_h = \psi_h^2$ is the square of the characteristic function of the symmetric migration distribution:

$$\psi_h = \frac{1}{2} \left(\sum_{j=0}^{n_{\mathrm{d}}-1} m_j e^{ijz(h)} + \sum_{j=0}^{n_{\mathrm{d}}-1} m_j e^{-ijz(h)} \right). \tag{A8}$$

This equation entails that individuals move clockwise on the circle (to deme $i = 1, 2, 3, \ldots$ on the right of the natal deme) and counterclockwise on the circle (to deme $i = -1, -2, -3, \ldots$ on the left of the natal deme) with identical distributions (for a development of isolation by distance models, see Rousset 2004, chap. 3).

The stationary probability of identity $Q_{k,t}$ between two individuals sampled at k steps apart on the lattice and at $t \geq 1$ generations apart is

$$Q_{k,t} = \frac{1 - Q_{0,0}}{Nn_{\mathrm{d}}} \sum_{h=0}^{n_{\mathrm{d}}-1} \frac{(\gamma\lambda_h)^{t/2}}{1 - \gamma\lambda_h} e^{-ikz(h)} \tag{A9}$$

(Malécot 1973, eq. [20]; Epperson 1999, eq. [B12]; Lehmann 2008, eq. [A33]). Note that for $t = 2$, equation (A9) is equal to equation (A7). This equality can intuitively be understood by considering the infinite island model of dispersal. In this case, a pair of individuals sampled in a deme in the same generation descend from the same ancestor with probability $(1 - m)^2/N$ (with probability $(1 - m)^2$, the two individuals are philopatric, in which case they descend from the same individual with probability $1/N$). A pair of individuals sampled at two generations apart from the same deme consists of both an ancestor and its descendant with the same probability $(1 - m)^2/N$ as above (with probability $(1 - m)^2$, the ancestral line of the individual sampled downstream was in the same deme two generations earlier, in which case it descends from the individual sampled in that generation with probability $1/N$).

The scalars $\lambda_h = \psi_h^2$ appearing in equations (A7) and (A9) are the eigenvalues of the backward migration matrix of a pair of genes, while the terms ψ_h are the eigenvalues of the migration matrix of a single gene. When there is a positive level of philopatry ($m_0 > 0$) and a line of descent of a gene is eventually able to reach every deme in the population in the long run, the migration matrix is regular (all the entries of some power of the matrix are positive) and it has a single unit eigenvalue given by $\psi_0 = \sum_{j=0}^{n_{\mathrm{d}}-1} m_j = 1$, while all other eigenvalues are strictly lower than 1 in absolute value: $0 \leq |\psi_h| < 1$ for $h \neq 0$ (a result that comes from the Perron-Frobenius theorem for nonnegative matrices; e.g., Nagylaki 1998, p. 1600).

Relatedness

Using the formulas for the probabilities of identity (eqq. [A7]–[A9]), the relatedness coefficient $R_{k,t}$ (eq. [2]) will be evaluated as a product of two limits:

$$R_{k,t} = \lim_{\mu \to 0} \frac{Q_{k,t} - \overline{Q}_t}{1 - Q_{0,0}} \times \lim_{\mu \to 0} \frac{1 - Q_{0,0}}{1 - \overline{Q}_0}. \tag{A10}$$

To this end, expand equation (A9) as

$$\begin{aligned} Q_{k,t} &= \frac{1 - Q_{0,0}}{Nn_{\mathrm{d}}} \left[\frac{(\gamma\lambda_0)^{t/2}}{1 - \gamma\lambda_0} e^{-ikz(0)} + \sum_{h=1}^{n_{\mathrm{d}}-1} \frac{(\gamma\lambda_h)^{t/2}}{1 - \gamma\lambda_h} e^{-ikz(h)} \right] \\ &= \frac{1 - Q_{0,0}}{Nn_{\mathrm{d}}} \frac{\gamma^{t/2}}{1 - \gamma} + \frac{1 - Q_{0,0}}{Nn_{\mathrm{d}}} \sum_{h=1}^{n_{\mathrm{d}}-1} \frac{(\gamma\lambda_h)^{t/2}}{1 - \gamma\lambda_h} e^{-ikz(h)}, \end{aligned} \tag{A11}$$

where the second line follows from using $\lambda_0 = 1$ and $e^{-ikz(0)} = e^{-ik0} = 1$. Inserting equation (A11) into the average probability of identity, one obtains

$$\begin{aligned} \overline{Q}_t &= \frac{1}{n_{\mathrm{d}}} \sum_{k=0}^{n_{\mathrm{d}}-1} Q_{k,t} \\ &= \frac{1}{n_{\mathrm{d}}} \sum_{k=0}^{n_{\mathrm{d}}-1} \left[\frac{1 - Q_{0,0}}{Nn_{\mathrm{d}}} \frac{\gamma^{t/2}}{1 - \gamma} + \frac{1 - Q_{0,0}}{Nn_{\mathrm{d}}} \sum_{h=1}^{n_{\mathrm{d}}-1} \frac{(\gamma\lambda_h)^{t/2}}{1 - \gamma\lambda_h} e^{-ikz(h)} \right] \\ &= \frac{1 - Q_{0,0}}{Nn_{\mathrm{d}}} \frac{\gamma^{t/2}}{1 - \gamma} + \frac{1 - Q_{0,0}}{Nn_{\mathrm{d}}} \sum_{h=1}^{n_{\mathrm{d}}-1} \frac{(\gamma\lambda_h)^{t/2}}{1 - \gamma\lambda_h} \frac{1}{n_{\mathrm{d}}} \sum_{k=0}^{n_{\mathrm{d}}-1} e^{-ikz(h)} \\ &= \frac{1 - Q_{0,0}}{Nn_{\mathrm{d}}} \frac{\gamma^{t/2}}{1 - \gamma}, \end{aligned} \tag{A12}$$

where the last line is obtained from the standard result that $\sum_{k=0}^{n_d-1} e^{-ikz(h)} = 0$ for all $h \neq 0$. Subtracting equation (A11) from equation (A12), one has

$$\lim_{\mu\to 0} \frac{Q_{k,t} - \overline{Q}_t}{1 - Q_{0,0}} = \frac{1}{Nn_d} \sum_{h=1}^{n_d-1} \frac{\lambda_h^{t/2}}{1 - \lambda_h} e^{-ikz(h)}. \tag{A13}$$

Using $\overline{Q}_0 = (1 - Q_{0,0})\gamma/[Nn_d(1 - \gamma)]$, where the right-hand side is obtained by setting $t = 2$ in the right-hand side of equation (A12) (or from eq. [3.70] in Rousset 2004), we have

$$\lim_{\mu\to 0} \frac{1 - Q_{0,0}}{1 - \overline{Q}_0} = \lim_{\mu\to 0} \frac{Nn_d(1 - \gamma)(1 - Q_{0,0})}{Nn_d(1 - \gamma) - \gamma(1 - Q_{0,0})}. \tag{A14}$$

It now remains to obtain an expression for $Q_{0,0}$, which can be done using equation (A7) and writing

$$\begin{aligned} Q_{0,0} &= \frac{1 - Q_{0,0}}{Nn_d}\left(\frac{\gamma}{1 - \gamma} + \sum_{h=1}^{n_d-1} \frac{\gamma\lambda_h}{1 - \gamma\lambda_h}\right) \\ &= \frac{1 - Q_{0,0}}{Nn_d}\left(\frac{\gamma}{1 - \gamma} + G(\gamma)\right), \end{aligned} \tag{A15}$$

where $G(\gamma) = \sum_{h=1}^{n_d-1} \gamma\lambda_h/(1 - \gamma\lambda_h)$, which gives

$$Q_{0,0} = \frac{\gamma + (1 - \gamma)G(\gamma)}{\gamma + (1 - \gamma)(Nn_d + G(\gamma))}. \tag{A16}$$

Substituting this equation into equation (A14) and simplifying yield

$$\lim_{\mu\to 0} \frac{1 - Q_{0,0}}{1 - \overline{Q}_0} = \frac{Nn_d}{Nn_d + G(1)}. \tag{A17}$$

Putting all terms together, the relatedness between two individuals sampled at k steps apart on the lattice and at $t = 1, 2, 3, \ldots$ generations apart can be expressed as

$$R_{k,t} = \frac{1}{Nn_d + \sum_{h=1}^{n_d-1} \lambda_h/(1 - \lambda_h)} \sum_{h=1}^{n_d-1} \frac{\lambda_h^{t/2}}{1 - \lambda_h} e^{-ikz(h)}, \tag{A18}$$

which is equation (2) with $\lambda_h = \psi_h^2$. Because $Q_{k,0} = Q_{k,2}$ (e.g., eqq. [A7]–[A9]), the relatedness coefficient between two individuals sampled at k steps apart in the same generation ($t = 0$) is obtained by substituting $t = 2$ into equation [A18]; that is, $R_{k,0} = R_{k,2}$.

Finally, in order to evaluate $R_{k,t}$ explicitly in terms of the migration distribution defined in the main text (e.g., figs. 1–4), I used equation (4) with equation (A8), which gives

$$\psi_h = (1 - m) + m \sum_{j=1}^{n_d-1} \frac{(1 - g)g^{j-1}}{1 - g^{n_d-1}} \cos(jz(h)). \tag{A19}$$