Ⓔ

The Relation between Multilocus Population Genetics and Social Evolution Theory

Andy Gardner,[1,2,*] Stuart A. West,[2,†] and Nicholas H. Barton[2,‡]

1. Department of Biology and Department of Mathematics and Statistics, Queen's University, Kingston, Ontario K7L 3N6, Canada;
2. Institute of Evolutionary Biology, University of Edinburgh, King's Buildings, Edinburgh EH9 3JT, United Kingdom

ABSTRACT: Evolution at multiple gene positions is complicated. Direct selection on one gene disturbs the evolutionary dynamics of associated genes. Recent years have seen the development of a multilocus methodology for modeling evolution at arbitrary numbers of gene positions with arbitrary dominance and epistatic relations, mode of inheritance, genetic linkage, and recombination. We show that the approach is conceptually analogous to social evolutionary methodology, which focuses on selection acting on associated individuals. In doing so, we (1) make explicit the links between the multilocus methodology and the foundations of social evolution theory, namely, Price's theorem and Hamilton's rule; (2) relate the multilocus approach to levels-of-selection and neighbor-modulated-fitness approaches in social evolution; (3) highlight the equivalence between genetical hitchhiking and kin selection; (4) demonstrate that the multilocus methodology allows for social evolutionary analyses involving coevolution of multiple traits and genetical associations between nonrelatives, including individuals of different species; (5) show that this methodology helps solve problems of dynamic sufficiency in social evolution theory; (6) form links between invasion criteria in multilocus systems and Hamilton's rule of kin selection; (7) illustrate the generality and exactness of Hamilton's rule, which has previously been described as an approximate, heuristic result.

Keywords: linkage disequilibrium, relatedness, Price equation, Hamilton's rule, kin selection, group selection.

* Corresponding author. Present address: St. John's College, Oxford OX1 3JP, United Kingdom; e-mail: andy.gardner@sjc.ox.ac.uk.

† E-mail: stu.west@ed.ac.uk.

‡ E-mail: n.barton@ed.ac.uk.

It has long been understood that describing the simultaneous evolutionary dynamics of multiple genes is difficult. Simplifying assumptions, such as statistical independence within and between loci (Hardy-Weinberg equilibrium and linkage equilibrium) and independence in fitness effects (additivity, multiplicativity), make analysis tractable but can give strongly misleading predictions. Allowing for associations between genes greatly enriches the evolutionary dynamics of populations, is crucial for our understanding of many evolutionary phenomena, and forms the basis of evolutionary modifier theory (Barton 2000). Typically, analytical exploration of multilocus models has been limited to two or three loci, with numerical simulations providing the only means of expanding to more complicated models.

An analytical methodology developed by Barton and Turelli (1991) and elaborated by Kirkpatrick et al. (2002) allows for arbitrary numbers of gene positions with arbitrary selective interactions and patterns of inheritance (see also Christiansen 1999). A general and flexible notation partitions the evolutionary process in a way similar to methods in classical genetics but makes explicit the various gene positions involved. It emphasizes the genetic associations over sets of gene positions, generalizing from the between-locus measure of linkage disequilibrium to associations among three or more genes at various positions within the individuals. It also implicitly allows for associations between genes in different individuals and thus provides a means for examining social evolutionary models (Kirkpatrick et al. 2002; Roze and Rousset 2005).

In this article, we highlight the social evolutionary interpretation of this multilocus methodology. In particular, we show how the quantitative genetical approach is exactly analogous to existing methodology used for social evolutionary problems and that it provides a straightforward guide to constructing and analyzing social evolutionary models of arbitrary complexity. Moreover, we demonstrate that multilocus theory may be used to solve problems of dynamic sufficiency in social evolution models so that we may examine coevolutionary dynamics of social evolu-

tionary traits and describe the evolution of relatedness itself. We emphasize that no new methodology is developed but rather that this article is intended as a synthesis of multilocus and social evolution theory so that the results derived within each of these bodies of theory may be readily interpreted in terms of the other.

We review multilocus and social evolution theory, establishing a link between the two using Price's (1970, 1972) theorem. Both bodies of theory partition total evolutionary change into selection and transmission components; we examine when these partitions are equivalent. In particular, we show that the multilocus approach can be applied to social evolution and that the results can be understood in terms of Hamilton's rule, the key concept in social evolution theory. We examine some simple illustrative models, involving the evolution of synergistic cooperation, the coevolution of cooperation and punishment, and the evolution of interspecific mutualism due to the buildup of genetic associations between species. These examples clarify the generality and mathematical validity of Hamilton's rule.

Price's Equation and the Genetical Theory of Natural Selection

The multilocus methodology and social evolution theory share a common root in what is essentially the quantitative genetic approach. The formal basis of all of evolutionary theory is population genetics, and the link between quantitative and population genetics is given by Price's theorem, which provides our starting point. In this section, we review Price's theorem as a mathematical identity underpinning a general theory of selection (Price 1972, 1995) and also in its more specific application to Fisher's (1930) genetical theory of natural selection. Price's equation partitions total evolutionary change into selection and transmission components. We discuss issues of dynamic sufficiency in relation to Price's approach.

Price's Theorem

Price's (1970, 1972) equation provides a formal basis for the theory of selection. It is a mathematical identity that fully describes evolutionary change. In its most general form, the equation emerges as a consequence of mapping individuals between generations and is straightforward to derive (Price 1972; Frank 1998). The equation focuses on a variable (z) and describes the change in its mean value over the course of a generation ($\Delta \bar{z}$) in terms of its association with individual relative fitness ($w/\bar{w}$) and the discrepancy between the values of an individual offspring and its parent ($\Delta z = z' - z$, where z' is the value of the variable in the offspring). Specifically, it states that

$$\Delta \bar{z} = \mathrm{Cov}\left(\frac{w}{\bar{w}}, z\right) + \mathrm{E}\left(\frac{w}{\bar{w}} \Delta z\right), \tag{1}$$

where "Cov" and "E" denote, respectively, covariance and expectation (arithmetic average) taken over all individuals in the parent population. Because of its minimal assumptions, this equation is of exceptional generality (Price 1972; Hamilton 1975; Frank 1995a, 1997b). Yet since the theorem follows directly from the definitions used, in a sense it does not tell us anything that we did not already know. The usefulness of this equation is that it neatly partitions the effects of selection between individuals (the covariance term) and transmission from parents to offspring (the expectation term) in a description of total change (Frank 1995a, 1997b, 1998).

While Price (1995; and see Frank 1995a) intended the theorem to underpin a general theory of selection that would apply to the selection of radio stations with the turning of a dial as readily as it describes evolutionary change (Price 1995), the form of Price's equation provided by Price (1970) is less general and was specifically intended to describe genetical evolution (Grafen 2006b). Here there is a mapping from parents to their successful gametes rather than to their offspring per se, and hence the approach more readily allows for sexual reproduction. Also, the variable z is intended to represent a unit of biological heredity. In general, it must be some linear function of genic values; for example, the frequency of a particular gene in the focal individual, the expected genic value of a gene drawn at random from the individual for a focal locus, or an arbitrary weighted sum of genic values across several loci (i.e., a "breeding value," or "p-score"; Price 1970; Falconer 1981; Grafen 1985a). This choice of variable means that under Mendelian inheritance—specifically, a fair meiosis—the change across a generation that is attributed to transmission effects is expected to be zero. This recovers Fisher's (1930) genetical theory of natural selection; selection is defined according to changes in gene frequencies. In particular, the statement that the change in the mean value of a trait is given by the covariance of that trait's breeding value and relative fitness has been referred to as the secondary theorem of natural selection (STNS; Robertson 1966, 1968), secondary to Fisher's fundamental theorem of natural selection (FTNS; Fisher 1930 [chap. 2], 1941; Edwards 1994). Together, these summarize the basis of Darwinism: those heritable traits positively associated with fitness will accumulate in biological populations (STNS), and hence the action of natural selection is to increase the mean fitness of the population (FTNS).

A difficulty is presented when populations contain different classes of individuals that make different contributions to future generations. How is a gene's frequency

to be averaged over such classes (Taylor 1990, 1996; Grafen 2006*b*)? We are not concerned with class-structured models in this article, yet it is instructive to briefly examine Fisher's solution to this problem, because it clarifies the formal theory of natural selection. Fisher's (1930, chap. 2) approach was to weight within-class gene frequency by the class's reproductive value (its relative asymptotic genetic contribution to future generations) when calculating the population average frequency (Price and Smith 1972; Taylor 1990, 1996; Grafen 2006*b*). To visualize Fisher's approach, consider a neutral gene that is "unlucky" enough to be found predominantly in low-reproductive-value classes. Its reproductive-value-weighted frequency across the population remains fixed even if, for the time being, its actual per capita frequency is decreasing. For a rare gene variant, the reproductive-value approach describes its actual asymptotic rate of increase in the context of a stable class distribution; that is, here the two measures of gene frequency coincide.

Dynamic Sufficiency

In applying Price's theorem to modeling of evolutionary problems, we encounter problems of dynamic sufficiency (Grafen 1985*a*; Queller 1992; Frank 1995*a*, 1997*b*, 1998). Prediction of the average value of a trait in the next generation requires more information than simply the average value of the trait in the present generation, and we need to know more than simply the average trait value in the next generation in order to calculate how this changes in subsequent generations. For example, even if we understand how trait value maps to relative fitness in subsequent generations, we do not know how the variance in trait values has changed between the parent and offspring generations, and therefore we cannot evaluate the covariance of trait value and relative fitness in future generations. We can determine the change in the variance using a further application of the Price equation, by treating the square of the trait value as the focal trait and seeing how the population average changes over the space of a generation. Yet this requires that we have information about yet higher moments of the trait distribution (Frank 1997*b*, 1998). In general, in order to give the approach dynamic sufficiency—that is, the ability to be applied recursively over an indefinite number of generations—higher-order moments must be expressed in terms of lower-order moments (moment closure; Barton and Turelli 1987; Frank 1995*a*, 1997*b*, 1998). Price's equation has been described as lacking dynamic sufficiency (e.g., Grafen 2000), yet the real issue is whether the model that it is applied to allows moment closure and hence is dynamically sufficient. Assumptions that provide dynamic sufficiency include the standard assumption of constant variance in quantitative-

genetic approaches and linkage equilibrium plus a finite number of alleles in multilocus models.

Multilocus Methodology

We now review the multilocus approach. We outline the need for a methodology that tracks genetic associations and not just gene frequencies, and we describe the approach developed by Barton and Turelli (1991) and Kirkpatrick et al. (2002). We present the general notation for describing populations and evolutionary processes and discuss how evolutionary change is partitioned into selection and transmission components within this framework. We then relate the multilocus mathematics to general principle of genetic hitchhiking. Finally, we discuss the concept of quasi-linkage equilibrium (QLE), a powerful simplifying feature of many multilocus analyses, and how this is attained through separation of timescales or, alternatively, an evolutionary-invasion approach. We relate the multilocus description of selection at QLE to Fisher's (1930) genetical theory of natural selection.

Why Have a Multilocus Methodology?

The formal basis of evolutionary theory rests upon population genetics, the study of mechanical population processes such as natural selection, mutation, migration, and random drift (Crow and Kimura 1970; Provine 1971). Proper prediction of the course of evolutionary change requires a full description of population composition at a given time step. The number of possible genotypes increases geometrically with the number of loci, and so multilocus analyses can be intractable. Typically, simulation is employed to examine models featuring three or more loci, although the results are typically opaque, and this is computationally expensive even for relatively few loci. A common simplification (e.g., Haldane 1964) is to assume statistical independence between loci (linkage equilibrium) so that the large number of genotype frequencies is determined by the smaller number of allele frequencies. Yet statistical associations between loci (linkage disequilibria) cannot, in general, be ignored, because these often arise through population processes, and it is well appreciated that indirect selection caused by direct selection on linked loci can dramatically alter the course of evolution. Such indirect "hitchhiking" effects are essential for understanding the evolution of sex and recombination (Maynard Smith 1978; Otto and Feldman 1997; Barton and Charlesworth 1998), the evolution of mate preferences (Kirkpatrick and Ryan 1991), gene flow through hybrid zones (Barton and Gale 1993), and how selection affects genetic diversity (Maynard Smith and Haigh 1974). In such cases,

it is essential that we follow the frequency of each genotype and not just gene frequencies.

Expressing evolutionary change in terms of changes in genotype frequencies can obscure the dynamics of quantities that are of more immediate interest, for example, gene frequencies and population average trait values. An alternative is to follow the gene frequencies and all linkage disequilibria, which is the approach adopted by the multilocus methodology of Barton and Turelli (1991) and Kirkpatrick et al. (2002). This involves tracking the same number of evolving variables as if we were following genotype frequencies, but it lends itself to a quantitative genetic approach that neatly partitions the various causes of evolutionary change. It also leads the way to a powerful simplifying assumption, that of quasi-linkage equilibrium (Kimura 1965; Nagylaki 1993), which reduces the multilocus problem to the same difficulty as is achieved by assuming linkage equilibrium but does approximate the effects of linkage disequilibrium.

Multilocus Notation

The power of the multilocus methodology lies in its generality, but this can make discussion of the interpretation of the notation somewhat confusing. We summarize the key notation in table 1. The following excursion into the notation will highlight only the features that are of most immediate interest to the aims of this article—that is, making the possibilities for modeling social evolution explicit. For a comprehensive account, see Kirkpatrick et al. (2002).

The approach is to describe the genetical composition of the population in terms of the allelic values at the various positions where genes reside and also the associations between these positions. For example, a model involving haploids with two biallelic loci might involve a separate position for each locus (and average allelic values, or allele frequencies, for each) and also a term describing the linkage disequilibrium between the two loci. But positions are not synonymous with loci—for instance, in a diploid context with genomic imprinting, it may be necessary to describe the maternal and paternal instances of the same locus as two separate positions. Thus, we might more correctly describe this as a multiposition methodology.

Of interest to us are the deviations ($\zeta_i = X_i - \wp_i$) of a gene's allelic value (X_i) from some arbitrary reference value ($\wp_i$) for a given gene position (**i**). It is natural (and convenient) to define the reference value as the average allelic value ($\wp_i = \overline{X}_i$) for that position so that the allelic deviations are simply deviations from the average. We use this convention throughout the remainder of this article. Deviation values may be assigned to sets of gene positions and are given by the product of the allelic deviations over the set ($\zeta_A = \prod_{i \in A} \zeta_i$). Statistical associations between gene positions are described by taking the expectation of these deviation terms across the population ($\mathfrak{D}_A = E(\zeta_A)$). For example, the association between two loci, i and j, in a haploid individual is $\mathfrak{D}_{ij} = E(\zeta_{ij}) = E[(X_i - \overline{X}_i)(X_j - \overline{X}_j)]$, which corresponds exactly to the linkage disequilibrium between these two loci. The definition allows us to describe the association between any two gene positions, such as two homologous genes at a single locus or an autosomal gene and a mitochondrial gene within a eukaryotic cell. It also allows us to generalize the association to arbitrary numbers of gene positions (fig. 1A). Because of our convention regarding reference values, the association at a single position is zero ($\mathfrak{D}_i = E(X_i - \overline{X}_i) = 0$).

Table 1: Summary of the key multilocus notation encountered in this article

| Multilocus notation	Definition
i, j	Generic gene positions
A, B	Generic sets of gene positions
X_i	Allelic value at a gene position **i**
$\wp_i$	Reference value for gene position **i**; usually the average, $\overline{X}_i$
$\zeta_i = X_i - \wp_i$	Allelic deviation at a gene position **i**
$\zeta_A = \prod_{i \in A} \zeta_i$	Allelic deviation for a set of positions **A**
$\mathfrak{D}_A = E(\zeta_A)$	Association over a set of positions **A**
p_i, q_i	Frequencies of the $X_i = 1$ and $X_i = 0$ alleles, respectively
z	Generic phenotype
G	Set of all gene positions contributing to the phenotype
γ_A	Contribution of set of positions **A** to the phenotype
w	Fitness
W	Set of all gene positions contributing to fitness
α_A	Contribution of set of positions **A** to relative fitness; the generalized selection coefficient
$t_{A \leftarrow U}$	Probability that set of positions **A** after transmission derived from set of positions **U** before transmission
T	Set of positions contributing to transmission phenotype
$\tau_{A \leftarrow U \mid V}$	Contribution of set of positions **V** to the $t_{A \leftarrow U}$ phenotype
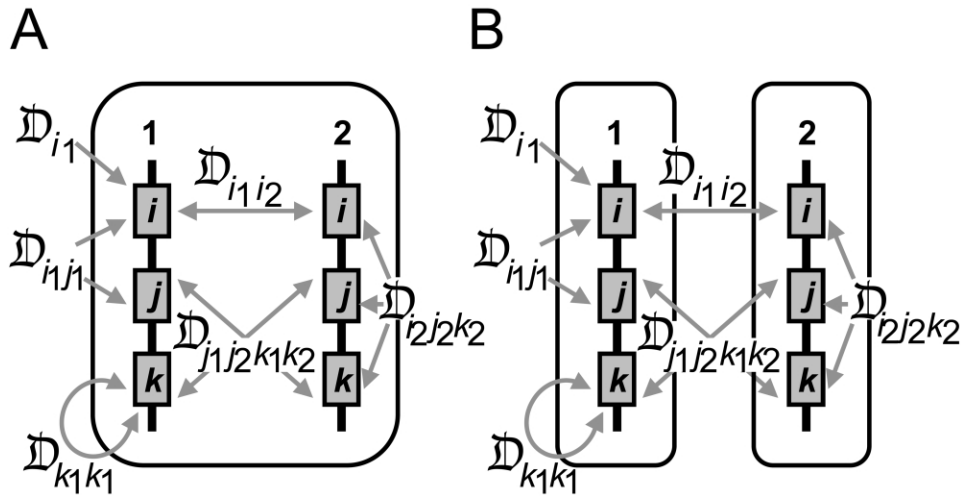
Figure 1: An illustration of gene positions and associations within the multilocus framework. *A*, A diploid individual contains one genomic set (*1*) from its mother and another genomic set (*2*) from its father. Here, three loci (*i*, *j*, and *k*) are shown. Association terms can be assigned to a single gene position (e.g., $\mathfrak{D}_{i_1}$), to two gene positions (e.g., linkage disequilibrium $\mathfrak{D}_{i_1j_1}$) within a genome set, the Hardy-Weinberg disequilibrium between maternal and paternal genomes at the same locus ($\mathfrak{D}_{i_1i_2}$), and the association of a gene position with itself ($\mathfrak{D}_{k_1k_1}$), and to three or more gene positions (e.g., $\mathfrak{D}_{i_2j_2k_2}$ or $\mathfrak{D}_{j_1j_2k_1k_2}$). *B*, There is no reason why associated genes must be restricted to the same individual. Here we illustrate some associations between two haploid individuals (*1*, *2*) with three loci. Note that this two-haploid-individuals model is mathematically equivalent to the diploid model.

Finally, there is no reason why associated genes should all be resident within a single individual (fig. 1*B*; Kirkpatrick et al. 2002; Roze and Rousset 2005). The major aim of this article is to expand on this crucial point and to forge conceptual links between multilocus and social evolution theory. Typically, we assume biallelic loci and assign one allele the value $X = 1$ and the other the value $X = 0$. Thus, the average allelic value at a gene position is the frequency of the $X = 1$ allele. We will denote this as p and the frequency of the $X = 0$ allele as $q = 1 - p$. This convention allows a reduction formula to be applied to association terms over sets containing repeated gene positions: $\mathfrak{D}_{\mathbf{Aii}} = p_iq_i\mathfrak{D}_{\mathbf{A}} + (1 - 2p_i)\mathfrak{D}_{\mathbf{Ai}}$. More generally, any number of alleles may exist at each locus, with arbitrary allelic values; the central multilocus expressions remain valid.

Once we have defined the genetical composition of the population in general, we can describe phenotypes. A phenotype (z) is defined as

$$z = \bar{z} + \sum_{\mathbf{A} \subseteq \mathbf{G}} \gamma_{\mathbf{A}}(\zeta_{\mathbf{A}} - \mathfrak{D}_{\mathbf{A}}), \qquad (2)$$

where $\bar{z}$ is the mean phenotypic value for the population, **G** is the set of all positions that contribute to the phenotype, and $\gamma_{\mathbf{A}}$ is the partial least squares regression of phenotype on the deviation term ($\zeta_{\mathbf{A}}$) for the set of loci **A** (holding all other deviations fixed). This is a standard multiple regression, where the predictors of phenotypic

value are given by the deviation terms $\zeta_{\mathbf{A}}$ (recall that the association $\mathfrak{D}_{\mathbf{A}}$ is simply the population average of $\zeta_{\mathbf{A}}$). We regard a phenotype in the broad sense, that is, any aspect (trait) associated with a set of gene positions, for example, weight, eye color, propensity to behave altruistically, or, perhaps, an allele's frequency or the association between genes in the set. A special case of particular interest is when the phenotype is relative fitness itself. We may express relative fitness as

$$\frac{w}{\bar{w}} = 1 + \sum_{\mathbf{A} \subseteq \mathbf{W}} \alpha_{\mathbf{A}}(\zeta_{\mathbf{A}} - \mathfrak{D}_{\mathbf{A}}). \qquad (3)$$

The $\alpha_{\mathbf{A}}$ terms provide generalized selection coefficients and are the partial regressions (i.e., holding all other associations constant) of relative fitness on the deviation ($\zeta_{\mathbf{A}}$) for a particular set of positions (**A**), and **W** is the set of all positions contributing to fitness. The relation of the generalized selection coefficient to average effect and average excess is discussed by Barton and Turelli (1991).

Describing Changes in Associations

The notation described above provides a sufficiently detailed description of population processes to allow for a dynamically sufficient model of multilocus evolutionary change, without simplifying assumptions such as linkage

equilibrium. Evolutionary change is summarized by two expressions describing selection and transmission in this system. The change in an association due to selection is

$$\Delta_S \mathfrak{D}_A = \sum_{U \subseteq W} \alpha_U (\mathfrak{D}_{UA} - \mathfrak{D}_U \mathfrak{D}_A) \qquad (4)$$

(Barton and Turelli 1991; Kirkpatrick et al. 2002). This description of change also allows us to describe the change in average allelic values at a single position. If we are concerned with biallelic loci with allelic values $X_i = 1$ at frequency p_i and $X_i = 0$ at frequency $q_i = 1 - p_i$, then from equation (4), the change in an allele's frequency is

$$\Delta_S p_i = \sum_{U \subseteq W} \alpha_U \mathfrak{D}_{Ui}, \qquad (5)$$

because $\Delta_S p_i = \Delta_S \mathfrak{D}_i$. Recall that Fisher's (1930) genetical theory of natural selection defines selection according to change in gene frequency, described in a peculiar way that weights by class reproductive values. This view of selection will not generally coincide with the description of selection provided by the multilocus approach, because the latter describes actual per capita gene frequency change. However, the two do coincide in the context of an evolutionary-invasion analysis, to be discussed below. The change in an association due to transmission is defined as

$$\Delta_T \mathfrak{D}_A = \left(\sum_U t_{A \leftarrow U} \mathfrak{D}_A \right) - \mathfrak{D}_A, \qquad (6)$$

where the $t_{A \leftarrow U}$ coefficients represent the probability that the set of positions $\mathbf{A}$ was drawn from source set of positions $\mathbf{U}$ during the transmission event. The above form is sufficient when transmission is genotype independent. Wherever this is not the case, for example, for segregation distorters (Hurst et al. 1996) or modifiers of recombination (Barton 1995), the transmission coefficient can be expressed as a regular phenotype:

$$t_{A \leftarrow U} = \bar{t}_{A \leftarrow U} + \sum_{V \subseteq T} \tau_{A \leftarrow U | V} (\zeta_V - \mathfrak{D}_V) \qquad (7)$$

(adapted from Barton 1995). Here, $\mathbf{T}$ is the set of all gene positions affecting the transmission phenotype, $\mathbf{V}$ is any combination of positions from this set, and the τ coefficients have the usual partial-regression definition. Genotype-dependent transmission is given by

$$\Delta_T \mathfrak{D}_A = \left\{ \sum_U \left[\bar{t}_{A \leftarrow U} \mathfrak{D}_U \right. \right.$$

$$\left. \left. + \sum_{V \subseteq T} \tau_{A \leftarrow U | V} (\mathfrak{D}_{UV} - \mathfrak{D}_U \mathfrak{D}_V) \right] \right\} - \mathfrak{D}_A. \qquad (8)$$

Note that, analogous to the derivation of expression (5) from equation (4), we can use expressions (6) and (8) to describe the change in allele frequency due to transmission. Allele frequencies do not change across genotype-independent transmission events, although they may change when transmission is genotype dependent.

Reference values ($\wp_i$) are not automatically updated during the selection or transmission event, and so if we used the average allelic values ($\overline{X}_i$) as reference values before the event, the associations after the event ($\mathfrak{D}_A'$) are still expressed in terms of deviations from the average allelic values before the event. In order to reexpress these in terms of deviations about the current average value, we need to update reference values, as outlined by Kirkpatrick et al. (2002). Kirkpatrick et al. also describe how deterministic population processes, such as mutation and migration, can be incorporated into the above scheme.

Hitchhiking

Such changes in gene frequencies and genetical associations can be described in terms of genetic "hitchhiking." This term was coined by Maynard Smith and Haigh (1974) to describe the increase in frequency for alleles that are linked to a beneficial mutation as the latter sweeps to fixation. It has since been used to describe, more generally, the perturbation of gene frequencies due to selection at linked loci (Barton 2000). An aim of this article is to include kin selection within this generalized view of hitchhiking. Here we show that the hitchhiking view allows for straightforward interpretation of the multilocus results.

Examining expression (5), which describes the change in a gene's frequency due to selection, and separating the response to direct selection at that gene position from indirect selection effects, we may write

$$\Delta_S p_i = \alpha_i \mathfrak{D}_{ii} + \sum_{U \subseteq W, U \neq i} \alpha_U \mathfrak{D}_{Ui}. \qquad (9)$$

The response to selection is given by the product of direct selection (α_i) acting on the focal gene position and the appropriate association term (the variance at that gene position, which for two alleles is $\mathfrak{D}_{ii} = p_i q_i$), plus the sum of similar products describing selection on other gene positions (α_U) and the association between these and the focal gene position ($\mathfrak{D}_{Ui}$). Assuming genotype-independent

transmission, this is the total change in gene frequency. For example, we may describe additive fitness effects s and t of genes at two linked loci i and j, respectively, within haploids. Here fitness is given by $w = 1 + sX_i + tX_j = 1 + sp_i + tp_j + s\zeta_i + t\zeta_j$. Using these to describe an average fitness function $w/\bar{w}$ and rearranging into the form of equation (3), we have generalized selection coefficients: $\alpha_i = s/\bar{w}$ and $\alpha_j = t/\bar{w}$. For this example, expression (9) takes the form

$$\Delta_s p_{\mathbf{i}} = \frac{s}{\bar{w}} p_i q_i + \frac{t}{\bar{w}} \mathcal{D}_{ij}, \qquad (10)$$

where the association modulating the effect of direct selection on locus j on gene frequency change at locus i is simply the linkage disequilibrium between these two loci $(\mathcal{D}_{ij})$. In the absence of indirect effects from the linked locus j, the direction of selection at i has the same sign as s; that is, the frequency of the focal gene at locus i increases when $s > 0$. If indirect effects are allowed for, this condition is $s + t\beta_{X_j, X_i} > 0$, where $\beta_{X_j, X_i} = \mathcal{D}_{ij}/p_i q_i$ is the regression of X_j on X_i. Thus, a neutral $(s = 0)$ or even deleterious $(s < 0)$ gene may increase in frequency under the action of selection at a linked locus $(t > 0$ and $s + t\beta_{X_j, X_i} > 0)$, and a beneficial gene $(s > 0)$ may decrease in frequency if its associate is sufficiently disfavored $(t < 0$ and $s + t\beta_{X_j, X_i} < 0)$.

Quasi Linkage Equilibrium

Kimura (1965) showed that multilocus systems often rapidly settle into a state such that linkage disequilibria (measured in a particular way, which allows them to be independent of gene frequencies) become virtually constant. He referred to this state as quasi linkage equilibrium (QLE; see also Nagylaki 1993). A separation-of-timescales approach based on this observation can greatly simplify a multilocus analysis. If the dynamics of linkage disequilibria occur on much shorter timescales than gene frequency change, then at any time the population can be fully described by the set of gene frequencies because knowing these allows recovery of the quasi-equilibrium genetic associations (Barton and Turelli 1991; Kirkpatrick et al. 2002). This separation of timescales is justified when processes that lead to the buildup of linkage disequilibrium (e.g., selection) are weak relative to those that lead to the breakup of linkage disequilibrium (e.g., recombination).

The QLE state is also attained when a very rare gene invades a population. The change in frequency of a vanishingly rare gene variant depends on its distribution across background genotypes, and although this distribution may change over time, it will eventually settle into some asymptotic state. In order to assess invasibility, one

considers the asymptotic rate of increase of the gene at this QLE state. The approach easily extends to allow for multiple loci with simultaneously segregating variation. In particular, the asymptotic growth of all gene variants is given by the leading eigenvalue and the stable distribution by the dominant right eigenvector of the transition matrix that describes how the population composition changes from one generation to the next. For a recent example of evolutionary-invasion analysis in a multilocus model that makes links with evolutionary game theory concepts such as evolutionary and convergence stability (Maynard Smith and Price 1973; Eshel and Motro 1981; Taylor 1996), see Gardner and Kalinka (2006). Evolutionary-invasion approaches are common in theoretical social evolution analyses and so provide a straightforward means of making tractable a social evolutionary multilocus analysis. The approach also allows one to sidestep issues of reproductive value in class-structured populations; recall Fisher's (1930) approach, in which reproductive-value weightings translate present gene frequency change into asymptotic change for a rare gene variant. Class structure is easily incorporated into multilocus analysis, simply by incorporating extra gene positions in order to distinguish genes in different classes of individuals (Kirkpatrick et al. 2002). Employing an evolutionary-invasion analysis to examine the asymptotic increase of a rare gene variant implicitly takes care of these reproductive-value issues. Put another way, the Fisherian account of natural selection and the description of the action of selection provided by the multilocus approach coincide in the context of an evolutionary-invasion analysis.

Foundations of Social Evolution Theory

We review the key concepts and formal foundations of social evolution theory. The link between alternative levels-of-selection and neighbor-modulated-fitness approaches to describing and modeling social evolution is given by Price's theorem. The neighbor-modulated-fitness approach gives rise to Hamilton's rule and the concept of relatedness. We discuss the inclusive fitness interpretation but suggest that a direct, neighbor-modulated-fitness approach is often simpler from a modeling perspective.

What Is Social Evolution Theory?

Social evolution concerns those traits that influence the fitness of individuals beyond their bearer (Trivers 1985; Frank 1998). These are categorized according to the fitness effects for the bearer of the trait and that individual's social partners. The four possibilities are mutual benefit $(+\ +)$, selfishness $(+\ -)$, altruism $(-\ +)$, and spite $(-\ -;$ table 2; Hamilton 1964; Trivers 1985; West et al. 2007). Typically,

Table 2: A classification of social behaviors according to their effect on the fitness of the actor and the recipient

Effect on actor	Effect on recipient	
	+	−
+	Mutual benefit	Selfishness
−	Altruism	Spite

Note: After Hamilton (1964), Trivers (1985), and West et al. 2007.

behaviors are described as cooperation, or helping, if they enhance a social partner's fitness (+ + and − +) and as harming if they reduce the social partner's fitness (+ − and − −; Rousset 2004; West et al. 2007). For a comprehensive synthesis of social evolution theory, see Frank (1998).

Levels of Selection versus Neighbor-Modulated Fitness

A formal basis for social evolution theory is provided by Price's (1970, 1972) theorem, which may be applied in two different but equivalent ways. Its first application is the levels-of-selection approach (Price 1972; Hamilton 1975; Wade 1985), which describes evolutionary change according to selection between groups and transmission of individuals between parent and offspring groups (incorporating a lower-level, within-group selection process). Making the between- and within-group selection processes explicit, Price's equation (1) becomes

$$\Delta \bar{z} = \text{Cov}_I\left(\frac{w_i}{\bar{w}}, z_i\right) + \text{E}_I\left[\text{Cov}_J\left(\frac{w_{ij}}{\bar{w}}, z_{ij} | i\right)\right], \quad (11)$$

where groups are indexed $i \in I$ and individuals within groups are indexed $j \in J$ so that z_{ij} is the trait value of the jth individual in the ith group, w_{ij} is this individual's fitness, $z_i = \text{E}_J(z_{ij})$ is the trait value of the ith group, $w_i = \text{E}_J(w_{ij})$ is the fitness of the ith group, and the subscripts I and J indicate the sets over which the summary statistics are taken.

The second application of Price's theorem to social evolution is the neighbor-modulated-fitness approach (Hamilton 1964, 1970). Here the same two-level selection process described by equation (11) is summarized into a single covariance term, where the target "neighbor-modulated" fitness summarizes the net effect of within- and between-group selection:

$$\Delta \bar{z} = \text{Cov}\left(\frac{w_{ij}}{\bar{w}}, z_{ij}\right). \quad (12)$$

Here the covariance between trait value and relative neighbor-modulated fitness is taken over all individuals in the whole population. This is simply the selection covariance component of the original Price (1970) equation. Note that we have explicitly assumed perfect transmission at the individual level. Transmission bias, for example, due to segregation distortion (Hurst et al. 1996), can be implemented within a neighbor-modulated-fitness scheme by considering these genes to be the individuals of the analysis. The neighbor-modulated-fitness and levels-of-selection approaches are mathematically equivalent. For example, in the context of the evolution of altruism, altruists suffer a within-group disadvantage ($\text{Cov}_J(w_{ij}/\bar{w}, z_{ij} | i) < 0$) due to exploitation by their more selfish social partners and enjoy a group-level advantage ($\text{Cov}_I(w_i/\bar{w}, z_i) > 0$) due to their altruism, which might under some conditions give a total advantage for altruism ($\text{Cov}(w_{ij}/\bar{w}, z_{ij}) > 0$). The neighbor-modulated fitness will reflect any tendency for altruistic individuals to associate with other altruists such that the benefits of socializing of altruistic neighbors might outweigh the immediate costs of altruism, to derive a net fitness benefit. This is the basis of the kin selection approach, with the association between social partners being described by the kin selection coefficient of relatedness. Thus, there is an equivalence between group selection and kin selection; they are the same process (Hamilton 1975; Wade 1985; Frank 1986b; Queller 1992; Dugatkin and Reeve 1994; Wenseleers et al. 2003; Foster et al. 2006). Although this means that the units-of-selection debate is empirically vacuous, it still manages to drag on in many of social evolution's sister disciplines (West et al. 2007). In practice, levels-of-selection approaches are cumbersome and technically problematic, and so the neighbor-modulated-fitness approach is often preferable (Frank 1986a, 1998; Taylor and Frank 1996; Queller 2004).

Hamilton's Rule

A central result of social evolution theory is Hamilton's (1963, 1964, 1970) rule, which is the basis of kin selection theory. The clearest and most general derivation of Hamilton's rule is achieved using Price's theorem (Price 1970; Hamilton 1970). Taking a neighbor-modulated-fitness approach and dropping the individual and group indices, the response to selection can be partitioned as a product of the selection gradient and additive genetic variation in the trait of interest:

$$\Delta \bar{z} = \text{Cov}\left(\frac{w}{\bar{w}}, z\right) = \beta_{w/\bar{w}, z}\sigma_z^2. \quad (13)$$

Assuming nonzero genetic variation, the response to se-

lection has the same sign as the regression coefficient, and this can be partitioned to give

$$\beta_{w/\hat{w},z} = \frac{\beta_{w,z\cdot\hat{z}} + \beta_{w,\hat{z}\cdot z}\beta_{\hat{z},z}}{\bar{w}} \tag{14}$$

(Queller 1992; Frank 1998), where the partial regression of the individual's fitness on its breeding value (holding fixed the average breeding value of social partners, $\hat{z}$) is $\beta_{w,z\cdot\hat{z}} = -C$, which describes the direct cost of carrying genes associated with the social trait; the partial regression of its fitness on social-partner breeding value (holding its own breeding value fixed) is $\beta_{w,\hat{z}\cdot z} = B$, which describes the benefit of having social partners genetically predisposed to the social behavior; and the regression of social-partner breeding value on its own breeding value is $\beta_{\hat{z},z}$, which is the kin selection coefficient of relatedness (R; Hamilton 1963, 1970; Grafen 1985*a*; Queller 1992) between the focal individual and its social partners. Substituting these definitions into expression (14), we derive a condition for when social trait is favored by selection:

$$RB - C > 0. \tag{15}$$

Hamilton's rule states that the trait is favored when the direct cost to the bearer is less than the direct benefit to social partners devalued according to the relatedness between bearer and recipient. The derivation reveals that Hamilton's rule is an exact and general statement of social evolutionary change. Here we have continued to neglect class structure; development of Hamilton's rule for class-structured populations is provided by Taylor (1990). For the purposes of this article, we simply note that this describes change in reproductive-value-weighted gene frequencies or, equivalently, the actual asymptotic gene frequency change.

Hamilton's rule provides a framework in which to conceptualize social evolution, but how is a social evolutionary analysis conducted? A straightforward approach is the direct-fitness methodology developed by Taylor and Frank (1996; see also Frank 1997*a*, 1998). This is a neighbor-modulated-fitness approach that assumes vanishing trait variation so that the statistical least squares regression of fitness on breeding value can be rewritten as a derivative of a fitness function with respect to genetic value, dw/dz. Analogous to the expansion of the regression into partial-regression terms in expression (14), we may write

$$\frac{dw}{dz} = \frac{\partial w}{\partial z} + \frac{\partial w}{\partial \hat{z}}\frac{d\hat{z}}{dz} = \frac{\partial w}{\partial z} + \frac{\partial w}{\partial \hat{z}}R, \tag{16}$$

where $\partial w/\partial z = -C$ describes the cost of the social act, $\partial w/\partial \hat{z} = B$ describes the benefit of associating with neigh-

bors expressing the social act, $d\hat{z}/dz = R$ is the kin selection coefficient of relatedness, and these derivatives are evaluated at $z = \hat{z} = \bar{z}$. Equilibrium points (z^*, satisfying $dw/dz|_{z=\hat{z}=\bar{z}=z^*} = 0$) may be assessed for evolutionary and convergence stability by inspection of second derivatives (Maynard Smith and Price 1973; Eshel and Motro 1981; Christiansen 1991; Taylor 1996). The approach is readily adapted to social evolution in class-structured populations (Taylor and Frank 1996; Frank 1998; Taylor et al. 2007).

In the above discussion, we have focused on the neighbor-modulated-fitness interpretation of Hamilton's rule, and this is most natural from a modeling perspective. However, Hamilton's (1964, 1970) usual interpretation was one of "inclusive fitness," and this has remained a popular concept. Neighbor-modulated fitness describes the fitness effect of a focal individual's behavior and the behaviors of its social partners on the focal individual's fitness, and properly this is averaged over all carriers of a variant gene. It provides a quantity that is maximized (a maximand) under the action of selection; however, it is the gene rather than the individual that is assigned a neighbor-modulated fitness, and so this leads to the view of the gene as maximizing agent (Dawkins 1976). Hamilton (1964, 1970, 1996) was interested in recovering a maximizing principle for individuals, in the tradition of Darwin (1859) and Fisher (1930), and developed inclusive fitness accordingly. Inclusive fitness attributes all the fitness consequences of a single actor's behavior for itself and all social partners, weighting each fitness component according to the relatedness of the recipient to the actor, and thus presents a fitness-accounting approach that is alternative but equivalent to neighbor-modulated fitness (Hamilton 1964; Queller 1992; Frank 1997*a*; Taylor et al. 2007). Crucially, a Darwinian individual has control over its inclusive fitness but not its neighbor-modulated fitness, and so inclusive fitness provides a maximand for social individuals (Grafen 2006*a*).

We now address some points relating to the kin selection coefficient of relatedness. There has been great confusion in the literature as to what "relatedness" is, for several reasons. First, different derivations of Hamilton's rule take as their focal trait a variety of different quantities, and so the exact definition of relatedness varies between these derivations. In the original derivation of Hamilton's rule (1964) and many that have followed (e.g., Hamilton 1970; Rousset 2004), the trait of interest (z) has been the genic value at a single gene position. Thus, the regression coefficient of relatedness has been described in terms of identity in state, specifically, departures of this from the population average (Rousset 2004). The interest in social evolution in structured populations, where social partners will tend to be genealogical kin (Hamilton 1964), has led to relatedness being described by Wright's F statistics (e.g., Hamilton 1970; Michod and Hamilton 1980; Frank 1998;

Rousset 2004). Elsewhere, Hamilton's rule has been derived to express the change in the social behavior phenotype (e.g., Grafen 1985*a*; Queller 1985, 1992; Frank 1998), often considered as a quantitative trait with many underlying gene positions contributing. Here the coefficient of relatedness is treated as a regression of some summary measure of the individual's genetic value for that trait—such as a breeding value (Queller 1992), p-score (Grafen 1985*a*), or gene frequency (Hamilton 1964, 1970)—or simply as the regression of partner phenotype on own phenotype value (Queller 1985).

Second, much interest in social evolution has focused on examples where the genetical similarity of social partners is due to coancestry. It has frequently been useful to describe relatedness in terms of identity by descent (IBD; Malécot 1948), which provides an approximation of identity in state for rare genetic variants (see Rousset 2004 for a recent review). The result is that the coefficient of relatedness has often been confused with genealogical relationship. It is unfortunate that in coining the phrase "kin selection," Maynard Smith (1964) inadvertently obscured a more general principle underlying the process he helped to popularize. Hamilton (1964) illustrated this with a thought experiment involving a "greenbeard" (Dawkins 1976) gene that causes its bearer to display a distinctive phenotypic marker and also to dispense altruism preferentially toward bearers of the marker, regardless of genealogical relationship. Because recipients of altruism carry the gene, there is a positive association (relatedness) between social partners, and hence greenbeard altruism may be favored (for empirical examples, see Keller and Ross 1998; Queller et al. 2003).

Such confusion has caused the importance of Hamilton's rule to be overlooked. The most widely known derivation of Hamilton's rule is the original (Hamilton 1964), where it emerged from analysis of a one-locus model. This has led to Hamilton's rule being regarded as an approximate, heuristic result based on a simplified model (Feldman and Cavalli-Sforza 1978; Bulmer 1994; Schwartz 2002). However, the derivation of Hamilton's rule using Price's theorem (Hamilton 1970) applies very generally. The cost of this generality is that it hides a lot of detail, and so naive application of Hamilton's rule may lead to mistakes. For this reason, it is easier to use standard population genetics, game theory, or other methodologies to derive a condition for when the social trait of interest is favored by selection and then use Hamilton's rule as an aid for conceptualizing this result (Taylor and Frank 1996; Frank 1998).

Social Evolutionary Multilocus Methodology

In this section we make explicit the links between multilocus population genetics and social evolution theory.

The key is simply to understand that genetic associations between individuals are equivalent to genetic associations within individuals (Kirkpatrick et al. 2002; Axelrod et al. 2004; Gardner and West 2004*a*; Billiard and Lenormand 2005; Roze and Rousset 2005) so that relatedness and linkage disequilibrium are united in a common framework and kin selection can be understood within a generalized view of genetic hitchhiking. We show that applying Price's (1972) generalized theorem to the multilocus notation recovers the central multilocus dynamical equations. In particular, the Price's-theorem and multilocus approaches partition evolutionary change into selection and transmission components in equivalent ways; this agreement is crucial for translating results between approaches. The capacity of the multilocus approach to describe social evolution is implicit, and so no new methodology is derived. Instead, the potential for modeling social evolution is illustrated using a simple model of cooperation. We show how the results of a social evolutionary multilocus analysis can be interpreted in terms of Hamilton's rule and demonstrate that Hamilton's rule is an exact and generally correct result even when applied to complicated models involving gene interaction.

Price's Theorem for Multiple Gene Positions

We described above how dynamic sufficiency in the context of a population genetic analysis requires that we be able to express higher-level associations in terms of lower-level associations (Barton and Turelli 1987; Frank 1998) and thus produce a closed system of Price's (1972) equations describing the total change of the system. This requires a model, or rather a set of parameter values, that fully describes the population in terms of selection and transmission processes. These parameters are described as part of the multilocus notation. We now demonstrate that the central multilocus expressions for change due to selection and transmission emerge as a result of applying the general form of Price's theorem to the multilocus notation. For the selection component of Price's (1972) theorem, we have

$$
\begin{aligned}
\Delta_S \mathcal{D}_A &= \mathrm{Cov}\left(\frac{w}{\bar{w}}, \zeta_A\right) \\
&= \sum_{U \subseteq W} \beta_{w/\bar{w}, \zeta_U \cdot \{\zeta_V : V \neq U, V \subseteq W\}} \mathrm{Cov}(\zeta_A, \zeta_U) \quad (17) \\
&= \sum_{U \subseteq W} \alpha_U (\mathcal{D}_{UA} - \mathcal{D}_U \mathcal{D}_A),
\end{aligned}
$$

which we obtain by substitution for $w/\bar{w}$ from equation (3) (Barton and Turelli 1991). The relative fitness ($w/\bar{w}$) appearing as the target of selection in the first line im-

plicitly describes the effect of all the genotypic determinants of fitness that are statistically associated with the focal trait (ζ_A). The partial regression in the second line describes the partial effect on relative fitness of the deviation term for the set of gene positions **U**, holding all other deviation terms fixed, and this provides the generalized selection coefficient (α_U) that appears in the third line. Note that here the "individuals" are arbitrary and may represent organisms or groups of organisms. Applying the multilocus notation to the transmission component of Price's (1972) theorem, under the assumption of genotype-independent transmission, gives

$$
\begin{aligned}
\Delta_T \mathfrak{D}_A &= E\left[\left(\frac{w}{\bar{w}}\right)\Delta\zeta_A\right] \\[2mm]
&= E\left\{\left[1 + \sum_{V \subseteq W} \alpha_V(\zeta_V - \mathfrak{D}_V)\right]\right. \\[2mm]
&\quad \left.\times \left[\sum_{U\,:\,U=A} t_{A\leftarrow U}\mathfrak{D}'_U - \zeta_A\right]\right\} \\[2mm]
&= \sum_{U\,:\,U=A} t_{A\leftarrow U}\mathfrak{D}'_U \qquad (18) \\[2mm]
&\quad - \left[\mathfrak{D}_A + \sum_{V \subseteq W} \alpha_V(\mathfrak{D}_{AV} - \mathfrak{D}_A\mathfrak{D}_V)\right] \\[2mm]
&= \sum_{U\,:\,U=A} t_{A\leftarrow U}\mathfrak{D}'_U - \mathfrak{D}'_A.
\end{aligned}
$$

Again, we have substituted in the multilocus description of relative fitness ($w/\bar{w}$) given in equation (3), and we describe the change between offspring and parent ($\Delta\zeta_A$) in terms of multilocus transmission details. Rearranging, we recover the multilocus expression (6) for change in an association due to transmission. Although the derivation is more complicated, it can be shown that considerations of genotype-dependent transmission agree with expression (8). Thus, Price's (1972) transmission scheme corresponds exactly with the multilocus transmission framework, although the latter makes explicit the redistribution of gene positions that is implicit in Price's simple notation. The selection/transmission partition employed by the multilocus methodology is the same as that of Price's (1972) theorem. Crucially, the multilocus selection/transmission partition is not equivalent to Price's (1970) formulation, which takes the Fisherian approach of tracking reproductive-value-weighted genetic change rather than actual per capita gene frequency change (although these are equivalent in an evolutionary-invasion analysis).

With these difference equations, the change in each of the association terms ($\mathfrak{D}_A$) can be determined, giving a complete description of the population in the next time step. They can be applied recursively to give a full compositional description of populations placed further in the future. At any stage, expression (6) (the genotype-to-phenotype map) can be applied to give a complete phenotypic description of the population. Thus, Price's approach has been made dynamically sufficient and explicitly describes the effect of each combination of gene positions; the result is the multilocus methodology.

The Evolution of Cooperation

We now apply the multilocus methodology to social evolution, using a simple model of cooperation (fig. 2*A*, 2*B*), to illustrate how this is handled in a way conceptually similar to standard multilocus models. We assume an infinite population of asexual haploid individuals that form pairs to play a one-shot, symmetrical game in each generation. Pairings are such that with probability *R* an individual is paired with its genetic clone and with probability $1 - R$ an individual is paired with a random member of the population. Here, *R* is a model parameter that happens to coincide with the coefficient of relatedness. In each game, an individual may either cooperate (C), in which case, her baseline fitness of 1 unit is decremented by *c* units and her partner's fitness is incremented by *b* units, or else she may defect (D), which affects neither her nor her social partner's fitness. Initially, we assume that fitness components combine additively, although this is relaxed later. The social behavior is controlled by a single biallelic locus, with the $X = 1$ allele at frequency *p* coding for cooperation and the $X = 0$ allele at frequency $q = 1 - p$ coding for defection. After the game, individuals produce a large number of offspring by parthenogenesis, in proportion to their fitness, which form the next generation of individuals. We may write a fitness function in terms of allelic values,

$$
w = 1 - cX_1 + bX_2, \qquad (19)
$$

where the subscripts denote oneself (1) and one's social partner (2). Following the usual procedure, we may extract generalized selection coefficients (α_1 and α_2) describing the direct effect of each set of gene positions and generate an expression for change due to selection:

$$
\begin{aligned}
p' &= p + \sum_A \alpha_A \mathfrak{D}_{A1} = p + \alpha_1 \mathfrak{D}_{11} + \alpha_2 \mathfrak{D}_{12} \\[2mm]
&= p + \alpha_1 pq + \alpha_2 \mathfrak{D}_{12}, \qquad (20)
\end{aligned}
$$

where we have used the reduction formula to write $\mathfrak{D}_{11} = pq$. The cooperation allele increases in frequency when $\alpha_1 pq + \alpha_2 \mathfrak{D}_{12} > 0$ or, equivalently, $\alpha_1 + \alpha_2 \beta_{X_2, X_1} >$
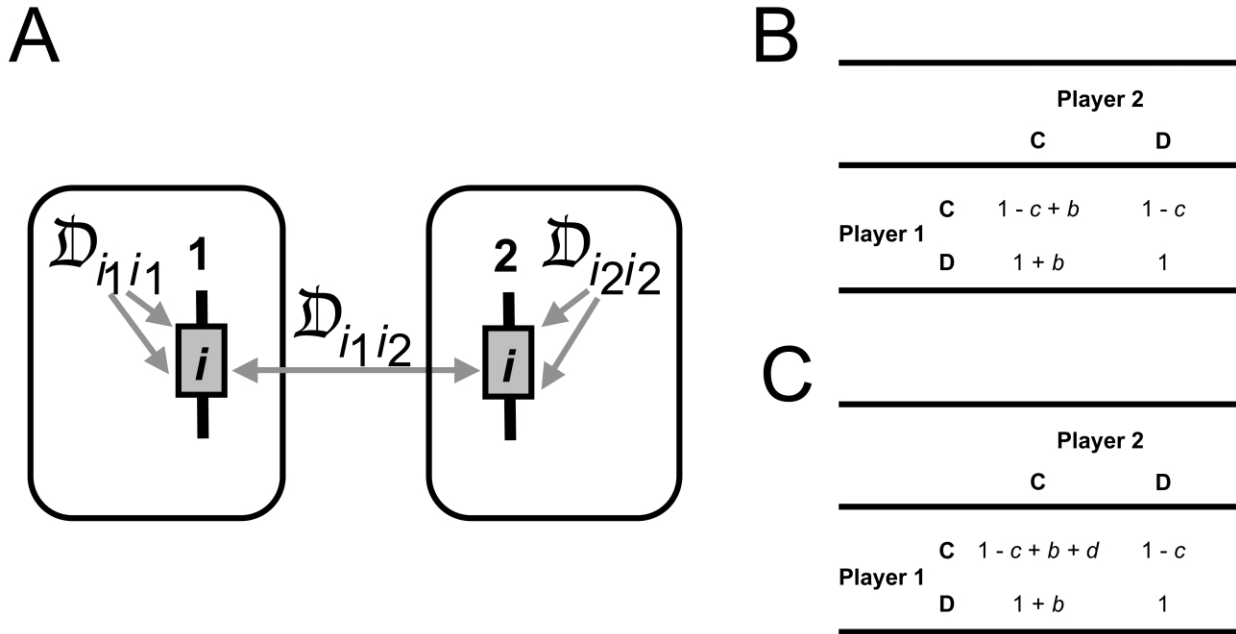
Figure 2: A model of kin-selected cooperation. *A,* A representation of the focal individual and her social partner; we consider a single locus and haploid genetics, so there are only two gene positions, *1* and *2*. The three associations that are of interest are highlighted: the variation at the focal gene position ($\mathcal{D}_{11} = pq$) and the variation at the associated gene position ($\mathcal{D}_{22} = pq$), plus the association between the two gene positions ($\mathcal{D}_{12} = Rpq$), where *R* is the kin selection coefficient of relatedness. *B,* Payoff matrix for a simple additive game: cooperation accrues a personal cost *c* and enhances the fitness of one's social partner by *b*. *C,* Payoff matrix for a synergistic game: in addition to the additive payoffs *b* and $-c$, cooperators accrue an extra synergistic payoff *d* in the company of other cooperators. *C* = cooperate; *D* = defect.

0, where $\beta_{X_2, X_1} = \mathcal{D}_{12}/pq$ is the regression of one's partner's allelic value on one's own allelic value (the kin selection coefficient of relatedness; the link with IBD in structured populations has been given by Roze and Rousset [2005]). Substituting the model parameters, the condition for increase is a Hamilton's rule:

$$Rb - c > 0. \tag{21}$$

Details of this derivation are given in appendix A in the online edition of the *American Naturalist.* Note that this condition is mathematically equivalent to that derived for the hitchhiking example presented earlier. As in the hitchhiking example, the net response to selection is given by the direct action of selection at the focal gene position plus the indirect action of selection spilling over from an associated gene position, and this indirect effect can allow for a directly deleterious gene to increase in frequency. Thus, kin selection can be seen as a special form of genetic hitchhiking.

Extending Hamilton's Rule

We assumed additive fitness components in the above model. Deviations from additivity are handled by the in-corporation of extra terms within Hamilton's rule. We illustrate this by extending the previous additive model so that when two cooperators interact, they accrue an extra increment (*d*) to their fitness (fig. 2*C*), which is given by

$$w = 1 - cX_1 + bX_2 + dX_1X_2. \tag{22}$$

The action of selection on the cooperation allele is described by

$$p' = p + \sum_A \alpha_A \mathcal{D}_{A1}$$
$$= p + \alpha_1 \mathcal{D}_{11} + \alpha_2 \mathcal{D}_{12} + \alpha_{12} \mathcal{D}_{112}. \tag{23}$$

Applying the reduction formula $\mathcal{D}_{Aii} = p_i q_i \mathcal{D}_A + (1 - 2p_i)\mathcal{D}_{Ai}$, we note that $\mathcal{D}_{11} = pq$ and $\mathcal{D}_{112} = (1 - 2p)\mathcal{D}_{12}$. The cooperation allele increases in frequency when $p' > p$, that is, when $\alpha_1 pq + \alpha_2 \mathcal{D}_{12} + \alpha_{12}(1 - 2p)\mathcal{D}_{12} > 0$. In terms of the model parameters, this is can be written as

$$R(b + dp) - c + dp + (1 - 2p)Rd > 0. \tag{24}$$

(A detailed derivation of the above is given in app. A.)

Queller (1985; see also Wenseleers 2006) explores an

equivalent model of kin selection with nonadditive fitness components, where the approach is essentially to regress fitness on the values X_1, X_2, and X_1X_2, which are treated as separate predictors. This approach generates the Hamilton's rule $Rb - c + [R + (1 - R)p]d > 0$, which is simply a rearrangement of equation (24). Queller's (1985) interpretation of that analysis is that the Hamilton's rule $Rb - c > 0$ is incorrect for nonadditive fitness components and that a new "synergy" coefficient ($S = R + (1 - R)p$), analogous to the coefficient of relatedness, is required to make Hamilton's rule "work." However, the inequality $Rb - c > 0$ is a naive statement of Hamilton's rule and in particular is not the rule derived by Hamilton (1970). Queller's (1985) synergy term disappears when benefits and costs are defined as the usual average effects on individual fitness rather than as the arbitrary model parameters b and c (Grafen 1985b). In particular, the condition for increase can be rearranged into the form $RB - C > 0$, where $B = b + \{[R + (1 - R)p]/(1 + R)\}d$ and $C = c - \{[R + (1 - R)p]/(1 + R)\}d$ are the proper fitness components of Hamilton's rule; that is, they satisfy their partial regression definitions ($B = \beta_{w,\hat{z} \cdot z}$ and $-C = \beta_{w,z \cdot \hat{z}}$, as described in "Hamilton's Rule"; see app. A for details). Thus, Hamilton's rule remains a correct statement (Grafen 1985b), albeit one in which the cost and benefit terms are somewhat complicated. This is not to say that Queller's (1985) partition is incorrect; indeed, it usefully makes the components of fitness more explicit and mechanistic. The multilocus approach makes explicit the direct action of selection on each set of gene positions, and it frames social evolutionary problems within a generalized view of genetic hitchhiking.

The model described here can be extended to arbitrary numbers of loci and individuals, with arbitrary ploidy, dominance, and epistatic and synergistic fitness effects, simply by applying the central multilocus equations. The next two sections illustrate the use of this approach for social evolutionary problems.

Example: Coevolution of Cooperation and Punishment

Explaining cooperation remains one of the greatest challenges for evolutionary theory (Maynard Smith and Szathmáry 1995; Hamilton 1996). Recent years have seen growing interest in the threat of punishment as an incentive for cooperation, particularly in relation to human behavior (Hirshleifer and Rasmusen 1989; Boyd and Richerson 1992; Sober and Wilson 1998; Sell and Wilson 1999). Empirical studies have repeatedly demonstrated a willingness in humans for investment in mutually costly punishing behaviors and that this sustains cooperation (Fehr and Gächter 2000, 2002). Yet the existence of costly punishment itself presents a problem and has stimulated some

theoretical investigation (Gintis 2000; Heinrich and Boyd 2001; Boyd et al. 2003; Panchanathan and Boyd 2004). One argument (e.g., Sober and Wilson 1998) suggests that punishment will often be cheaper than cooperation so that kin selection can maintain punishment even when relatedness is too low for cooperation to be directly favored. Hence, kin selection maintains punishment, and punishment maintains cooperation. This explanation has received some mathematical attention (Gardner and West 2004a) and has been rejected on conceptual grounds. Since punishment directly harms both punisher and punished, increased relatedness between social partners directly disfavors the evolution of punishment. Gardner and West (2004a) showed that a different association between individuals—the association between the cooperation strategy of one's social partners and one's own punishment strategy—is crucial for the evolution of punishment. Yet the analysis suffered from problems of dynamic insufficiency and, in particular, could not predict the dynamics of this key association. We apply the multilocus methodology to determine whether and when punishment and cooperation will evolve within a simple illustrative model.

We restrict attention to haploid individuals interacting in pairs in one-shot games. For simplicity, we consider that social partners are genetically identical with (fixed) probability R and are randomly pairing with probability $1 - R$. One individual per pair is randomly assigned the role of player 1, and the other is player 2. Player 1 has the option of cooperating (C), incurring a cost c for itself and a benefit b for player 2, or defecting (D), with no direct consequences for fitness. If player 1 defects, then player 2 responds by either punishing (P), incurring a cost a for itself and a cost d for player 1, or forgiving (F), which has no direct fitness consequences. We assume a, b, c, $d > 0$. Each player's cooperation strategy is controlled by a single biallelic locus, i, with allelic values $X_i = 1$ for cooperation and $X_i = 0$ for defection. Punishment strategy is controlled by a single biallelic locus, j, with allelic values $X_j = 1$ for punishing and $X_j = 0$ for forgiving. To allow for inherent costs of the punishing strategy, for example, resource expenditure for monitoring partner play, all individuals adopting the punishing strategy will also incur a cost $e \geq 0$ when playing the role of player 2, regardless of player 1's strategy. Individuals produce a number of successful clonal gametes proportional to their payoff. Gametes pair to form diploid zygotes, which then undergo meiosis to form the new generation of haploid individuals. The effective recombination rate between loci i and j, describing the effects of physical linkage and inbreeding, is denoted r. Genotypic fitness can be written in the form

$$w = 1 - \frac{1}{2}\big[cX_{i_1} + d(1 - X_{i_1})X_{j_2}\big]$$

$$+ \frac{1}{2}\big\{X_{i_2}b - X_{j_1}\big[e + (1 - X_{i_2})a\big]\big\}, \quad (25)$$

where the numeric subscripts indicate whether the gene position is within the focal individual (1) or the social partner (2). We then follow the usual procedure to determine the dynamics of the allele frequencies and linkage disequilibrium (see app. B in the online edition of the *American Naturalist*). Having derived the necessary recursions, we determine invasion and maintenance conditions.

When can cooperation and punishment invade? In other words, when is a defecting-forgiving population not an evolutionarily stable state? To examine this, we consider that both cooperation and punishment alleles are vanishingly rare in the population and examine their asymptotic rate of increase. We find that cooperation and punishment cannot invade unless $Rb - c > 0$ (details in app. B), which is the condition for cooperation to be favored in the absence of punishment. Since the interest in punishment is due to its potential for allowing cooperation where this is otherwise difficult to explain, we conclude that cooperation and punishment cannot invade in any interesting situations.

When is cooperating and punishing evolutionarily stable? To answer this, we examine the invasion of vanishingly rare defection and forgiveness alleles (details in app. B). We find that when punishment has an intrinsic cost ($e > 0$), then cooperation-punishment is not an evolutionarily stable population state. Allowing for zero intrinsic cost of punishing ($e = 0$), a population fixed for cooperation and punishment alleles is only neutrally stable, and so it can always be invaded by the more forgiving allele by neutral drift. Therefore, we conclude that cooperation and punishment are not evolutionarily stable within this simple model.

Although these results are consistent with the open-model analysis of Gardner and West (2004*a*), they are also different in important ways. Gardner and West (2004*a*) showed that in such a coevolutionary model, increasing relatedness (R) between social partners directly selects against punishment, because this is a mutually detrimental act that derives personal and kin-selected costs. Rather, the only benefit of punishment is the possibility for punishers to tend to associate with cooperators, this association being the product of a within-locus, between-individual association (relatedness, R) and a between-locus, within-individual association (linkage disequilibrium, $\mathcal{D}_{ij}$). Thus, while relatedness directly acts to disfavor punishment, it is required in order for punishment to accrue this indirect benefit. The open-model approach of Gardner and West (2004*a*) was unable to quantify these two opposing effects. The closed multi-locus approach taken here has shown that the direct negative effect will always overpower the indirect positive effect, giving a net disadvantage for punishment.

For simplicity, we have assumed that the relatedness of social partners is a fixed parameter. More generally, this may be an evolving variable. Roze and Rousset (2005) present a neighbor-modulated-fitness approach that can be extended to examine changes in these associations. We illustrate a levels-of-selection approach to this problem in the next section.

Example: The Evolution of Relatedness

An underexamined, though potentially important, effect in social evolution is the ability of selection to generate associations between individuals (Frank 1994; Pepper and Smuts 2002; Ajar 2003; Roze and Rousset 2005; Foster and Wenseleers 2006). Frank (1994) has shown in a model of interspecific mutualism that group selection favoring cooperative groups will cause cooperators to tend to associate with cooperators of the other species. Other studies (Frank 1995*c*; Foster and Wenseleers 2006) have demonstrated that such spontaneous emergence of between-species associations can allow for the evolution of between-species altruism, although this has relied on either numerical simulation or analytical results based on standard approaches that lack dynamic sufficiency. Thus, the process remains obscure. The multilocus approach provides a means of exploring this neglected phenomenon.

We consider a very simple model for illustration. Social groups consist of pairs of individuals, with one of each pair being drawn from a species 1 and the other from a species 2. Individuals are haploid and contain only one locus of interest, i, which controls cooperation strategy: bearers of the $X_i = 1$ allele are cooperators (C), and bearers of the $X_i = 0$ allele are defectors (D). For simplicity, we assume that the cooperator allele is at the same frequency, p, in both species. Social behaviors affect group fitness. A group consisting of two defectors (DD) gives rise to one daughter group, on average. A group comprising a cooperator and a defector gives rise to $1 + s$ daughter groups, on average. A group consisting of two cooperators (CC) gives rise to $(1 + s)^2 + \varepsilon$ daughter groups, on average. The deviation from a multiplicative fitness scheme is described by the synergy parameter ε. Social behaviors also affect individual fitness. A cooperator paired with a cooperator enjoys fitness $(1 + s)^2 + \varepsilon$, but a cooperator paired with a defector enjoys fitness $1 + s - f$. A defector paired with a defector enjoys fitness 1, and a defector paired with a cooperator enjoys fitness $1 + s + f$. Thus, cooperators suffer a within-group disadvantage, and defectors enjoy a within-group advantage in CD and DC

groups. The extent of this disparity is controlled by the parameter f. Note that a group's fitness is simply the average fitness of its constituent individuals. The above assumptions imply that DC and CD groups produce an excess of defectors of one species and a deficit of cooperators of the other species. We assume that this is compensated for by movement of excess defectors from DC groups into CD groups and vice versa, where they take the place of the missing cooperators and hence transform the group into a DD group (this assumption crucially requires that the frequency of C is the same in both species). After the new groups are established, groups dissociate into their constituent individuals with probability m, and the dissociated individuals re-pair at random.

A group's expected fitness may described as a function of its allelic values,

$$
\begin{aligned}
w = {} & X_1 X_2 [(1 + s)^2 + \varepsilon] \\
& + [X_1(1 - X_2) + (1 - X_1)X_2](1 + s) \\
& + (1 - X_1)(1 - X_2),
\end{aligned} \tag{26}
$$

where the subscripts denote the species in which the gene position occurs. We are also required to specify selection of individuals within groups, which is described in terms of transmission between parent and offspring groups. It is simpler to consider the genotype-dependent (individual selection) and genotype-independent (individual dispersal) transmission events separately. With regard to genotype-dependent transmission, we note that offspring groups are exact copies of parent groups unless the parent group contains a cooperator and a defector, in which case, with probability $f/(1 + s)$, the offspring group contains two defectors. Because of the symmetry, this can be described by four transmission functions. Denote one individual per group i (i.e., $i \in \{1, 2\}$) and the other j (i.e., $j \in \{1, 2\}$, $j \neq i$). Then the source of individual i in the daughter group is individual i in the parent group unless $X_i = 1$ and $X_j = 0$, in which case, in f out of the $1 + s$ daughter groups, the source is a foreign individual k of the same species as i, and $X_k = 0$ because the individual k is always a defector. The transmission functions are

$$
t_{i \leftarrow i} = 1 - X_i(1 - X_j)\frac{f}{1 + s},
$$

$$
t_{i \leftarrow k} = X_i(1 - X_j)\frac{f}{1 + s}, \tag{27a}
$$

$$
t_{\{i,j\} \leftarrow \{i,j\}} = 1 - [X_i(1 - X_j) + (1 - X_i)X_j]\frac{f}{1 + s},
$$

$$
t_{\{i,j\} \leftarrow \{k,j\}} = X_i(1 - X_j)\frac{f}{1 + s}. \tag{27b}
$$

From these may be extracted the various transmission coefficients required for the genotype-dependent transmission, expression (8), as described in appendix C in the online edition of the *American Naturalist*. We consider vanishingly rare cooperation in the population and determine the asymptotic rate of increase of the cooperation allele (full details of the analysis are given in app. C). We find that a sufficient condition for cooperation to increase from rarity is $f < s$. This corresponds to when cooperation is selfish, such that the fitness of a cooperator paired with a defector is greater than that of a defector paired with a defector ($1 + s - f > 1$), and this is therefore a trivial result. Restricting attention to the case where $f > s$, we find that cooperation is favored when $(1 - m)[(1 + s)^2 + \varepsilon] > 1$, which is of the form $1 + RB - C > 1$. When cooperation is favored, the between-species relatedness at quasi equilibrium is found to be $R = 1 - m[1 + (1 + s - f)/(s + s^2 + \varepsilon + f)]$. Thus, increasing migration of individuals (m) reduces relatedness, and increasing within-group conflict (f) and synergy (ε) increases relatedness, while the relation between the group benefit of cooperation (s) and relatedness is more complicated. The fitness components in Hamilton's rule are $B = s + f + (s^2 + \varepsilon)R/(1 - R)$ and $C = -s + f - (s^2 + \varepsilon)R/(1 - R)$. While the within-group conflict parameter features in each component of Hamilton's rule (R, B, and C are all functions of f), it cancels from Hamilton's rule itself ($RB - C$ is not a function of f). Increasing f reduces the success of cooperators partnered with defectors, disfavoring cooperation, but also increases the association between social partners (R), favoring cooperation, and these two effects exactly cancel in this simple model.

Here we have shown that the multilocus methodology may be applied to examine the evolution of relatedness. We have shown that genetical associations can emerge between social partners even when coancestry is ruled out, for example, where the social partners belong to different species. This relatedness settles to some asymptotic state, and this can be expressed analytically as a function of model parameters. Finally, we have shown that this relatedness can allow for altruistic cooperation between species. The condition for when cooperation will evolve is readily interpreted as a Hamilton's rule. The assumption of symmetry is made for convenience of analysis and will rarely be satisfied, because interspecific mutualisms often involve

disparate species with different (complementary) ways of life and hence different costs and benefits relating to the mutualism. Thus, while the above model is useful for illustration of general principles and the way in which the multilocus methodology can be used, this is an important evolutionary problem that deserves to be examined properly.

Discussion

Social Evolutionary Multilocus Methodology

We have formed links between multilocus population genetics and social evolution theory. This provides a general framework that can be used to model evolution at arbitrary numbers of gene positions across arbitrary numbers of individuals, with arbitrary effects on individual and social-partner fitnesses, with arbitrary dominance, epistasis, and synergy and arbitrary physical chromosomal linkage, mode of inheritance, and group dynamics. We emphasize that no new methodology has been developed. Rather, the effects of multiple gene positions are already implicitly accounted for in Price's theorem and Hamilton's rule, and the social evolutionary capabilities of the multilocus methodology are similarly implicit within the general and flexible notational scheme. We have made these hidden details explicit and have provided a simple, general framework in which to conduct analyses of social evolution at multiple gene positions.

   Multilocus models of social evolution have been constructed previously, using a range of methodologies. Axelrod et al. (2004) modeled tag-mediated altruism, using an extension of Price's (1970, 1972) selection-covariance mathematics. In their model, relatedness at an altruism locus is inferred by phenotypic similarity that indicates shared genetics at linked loci. Queller (1984, 1985) and Wenseleers (2001, 2006) used similar approaches to examine synergistic fitness interactions in a kin selection analysis incorporating multiple gene positions. The multilocus methodology provides a single general selection expression (8) that can be used to generate such selection-covariance expressions for models of arbitrary complexity. The transmission equation (10) also guides the modeler toward a complete set of closed equations that allow for dynamic sufficiency. More explicit applications of multilocus techniques to social evolution have been given by Billiard and Lenormand (2005) and especially by Roze and Rousset (2005), both in the context of the evolution of dispersal. In those analyses, the generalized view of linkage disequilibrium has been expressed in terms of probabilities of identity by descent (Malécot 1948) in the context of structured populations. A QLE approach, based on a separation of timescales, allowed the analyses to be simplified

to the level of gene frequencies. We have shown more broadly the connection between the multilocus and social evolution approaches stemming from their shared approach of partitioning fitness effects of genetic determinants in an essentially quantitative-genetic way.

   It is useful to compare the multilocus methodology that we have developed here with the direct- (neighbor-modulated) fitness approach of Taylor and Frank (1996) and Frank (1998; see also Rousset 2004). The latter derives from the standard evolutionary game theoretic approach of using differentiation techniques to identify evolutionary equilibria and then to characterize these in terms of evolutionary and convergence stability (Maynard Smith and Price 1973; Eshel and Motro 1981; Christiansen 1991; Taylor 1996). This approach avoids the conceptual minefield of constructing inclusive fitness functions and has greatly facilitated the simple and transparent analysis of many social evolutionary problems, including cooperation (Frank 1998; West and Buckling 2003; Gardner and West 2004a), cooperative breeding (Pen and Weissing 2000), sex allocation (Frank 1998), dispersal (Roze and Rousset 2005), virulence (Frank 1996a; Brown 1999; Gandon and Michalakis 2000; West and Buckling 2003), policing (Frank 1995b, 1996b, 2003; Wenseleers et al. 2004; Dionisio and Gordo 2006), sanctions and punishment (West et al. 2002; Gardner and West 2004a), spite (Gardner et al. 2004, 2007; Gardner and West 2004c), and mimicry (Johnstone 2002). The usual assumption is of vanishing variation around the current resident breeding value for each of the traits considered in the model so that there is effectively only one gene position that is undergoing evolutionary change at any particular time. The multilocus approach is of use when multiple traits are concerned or when a single focal trait has complex genetic architecture and where gene variants are segregating at several gene positions simultaneously. The direct-fitness approach of Taylor and Frank (1996) is readily extensible to the coevolution of multiple traits, such as punishment or policing and cooperation, resulting in the emergence of derivatives describing relatedness, linkage disequilibrium, and between-individual, between-trait associations (Gardner and West 2004a; Foster and Wenseleers 2006). However, the approach does not describe the evolution of such associations, so it is usual to assume statistical independence or a fixed association for coevolving traits in order to recover dynamic sufficiency (Frank 1995b, 1996b, 2003; Wenseleers et al. 2004; Gardner and West 2004a). The multilocus approach provides a simple, general notation and recursions for transmission of gene positions, allowing for the closure of such models, and hence the evolutionary dynamics of the genetic associations may be followed. This is important in models of cooperation and punishment and of the emergence of genetical associations between species, as we have

found in the course of this article. Thus, it is complementary to the direct-fitness approach. The social evolutionary multilocus approach would be useful in models of tag-mediated cooperation (e.g., Axelrod et al. 2004), whereby phenotypic similarity can be used to infer genetical similarity at the loci controlling cooperation.

During the development of social evolution theory, two equivalent approaches have been identified. The levels-of-selection approach distinguishes a response to between-group selection and a response to within-group selection. This may be extended to any number of nested levels, with each entity at each level being assigned a fitness value (Price 1972; Hamilton 1975). The neighbor-modulated-fitness approach focuses on the lowest level and summarizes the effects of selection at all levels within a single fitness term at this lowest level (Frank 1998). We have found the neighbor-modulated-fitness approach to be the easiest way to proceed with a multilocus analysis of social evolution because it avoids the complications of conducting a genotype-dependent transmission analysis at each level. This mirrors developments in the field of social evolution theory over the past decade, in which direct-fitness approaches have emerged as the most straightforward means of analysis (Taylor and Frank 1996; Frank 1998; Queller 2004; Rousset 2004). We have used the levels-of-selection approach to examine the evolution of associations between individuals, including between species. Doing this makes it immediately apparent how tedious and difficult such an approach is. In our simple two-gene model of interspecific mutualism, we were required to describe 16 separate transmission coefficients. Although these are readily generated using a simple algorithm, the four transmission functions from which they were derived are not so easily constructed. The number of required transmission functions will increase geometrically with the number of gene positions considered. Alternative approaches are possible, for example, ones based on a neighbor-modulated-fitness view that attributes a separate fitness function to every combination of sets of individuals (Roze and Rousset 2005; F. Rousset, personal communication), yet these are equivalent and do not simplify the problem. A major task for the future will be to develop simple general approaches that allow for the tracking of between-individual associations.

Relatedness and Hamilton's Rule

The multilocus approach provides not only an analytical tool for solving specific models but also a conceptual aid. By describing all genetical associations and resulting hitchhiking effects in a single intuitive framework, the multilocus methodology highlights conceptual equivalences that facilitate general understanding and may promote novel insight. Of particular interest in this article are the relation between relatedness and linkage disequilibrium and the associated equivalence of models of kin selection and models of genetical hitchhiking. What is crucial for hitchhiking is a statistical association between gene positions, such that selection at one position spills over onto the other (Barton 2000). This may be due to loci being situated on the same chromosome, but physical linkage is not strictly required. Similarly, models of kin selection (broadly defined) involve genetical associations between individuals, described as "relatedness." This may be due to close genealogical relationship between social partners, although, again, this is not a necessity (Hamilton 1964; Grafen 1985*a*; Nee 1989; Frank 1995*a*, 1998). Natural selection operates on genes with respect to their correlation with relative fitness and is blind to the causal mechanisms underlying this correlation. Previously, this conceptual link has been obscured because of terminology that emphasizes only certain causes of such associations—relatedness implying coancestry and linkage disequilibrium implying physical linkage. The multilocus approach may help to dispel misconceptions regarding relatedness by demonstrating the conceptual link with linkage disequilibrium. For example, while it is well accepted that linkage disequilibrium can take negative values when there is a tendency for the focal genes to exist apart from one another, the possibility of nontrivial negative relatedness has remained controversial (Foster et al. 2001; Gardner and West 2004*b*, 2004*c*). Negative relatedness arises when social partners are less alike than average (Grafen 1985*a*; Frank 1995*a*; Pepper 2000) and is implicated in the evolution of spiteful behaviors (Hamilton 1970; Grafen 1985*a*; Foster et al. 2001; Gardner and West 2004*b*, 2004*c*).

We have shown that the selection/transmission partition employed by the multilocus methodology is the same as that adopted by Price (1972), and hence we may relate the results of multilocus analyses to the body of social evolution theory that is derived from Price's theorem, in which Hamilton's (1970) rule is of central importance. However, there is a subtlety here, as a result of Hamilton's work being built on the foundations laid by Fisher, in which the action of selection is described in terms of changes in a peculiar reproductive-value-weighted measure of gene frequency (Fisher 1930 [chap. 2], 1941; Price 1970; Price and Smith 1972; Taylor 1990). In contrast, the multilocus methodology describes the action of selection in terms of actual per capita gene frequency change. Fortunately, these two descriptions of gene frequency change are equivalent for a rare gene variant that has attained its asymptotic state in terms of a stable class distribution and QLE, so that the results of evolutionary-invasion analyses that employ the multilocus approach may readily be interpreted in terms of Hamilton's rule.

A central message of this article is that the foundations of social evolution theory are solid and encompass models of arbitrary complexity. Price's theorem and Hamilton's rule allow for the effects of interacting gene positions, although the cost for this generality within such compact expressions is that they are rather nebulous and subtle statements. This has led to their being misunderstood and misused. For example, (1) it has been suggested that social evolution theory lacks an underlying general methodology that would allow models from different areas, such as intragenomic conflict and cooperative breeding vertebrates and interspecific mutualisms, to be described in the same terms (Herre 1999; Charlesworth 2000), and (2) it is generally regarded that Hamilton's rule is a heuristic result that works only under the assumption of weak selection (Charlesworth 1978; Feldman and Cavalli-Sforza 1978; Bulmer 1994; Wenseleers 2001, 2006; Schwartz 2002). As we have shown, Hamilton's rule is of such generality that it remains valid for multilocus models, models of interspecific mutualism, and arbitrary strength of selection.

However, a crucial point here is that the generality of Hamilton's rule is possible only because of the subtlety of its component terms, in which potentially complicated details are implicit. The same caveat applies to Price's theorem, from which Hamilton's rule has been derived. This means that Hamilton's rule and Price's theorem should generally be used in the interpretation of theory and not as the starting points in the analysis of specific problems, because this can easily lead to mistakes (Price 1970, 1972; Taylor and Frank 1996; Frank 1998; Pen and Weissing 2000; Gardner and West 2004c). The most powerful and simple approach to social evolutionary problems is to start with a method such as population genetics (including the multilocus approach), game theory, or direct-fitness maximization techniques. The results of these analyses can then be interpreted within the frameworks that Price's theorem and Hamilton's rule provide. The correct use of these powerful theorems is to translate the results of such disparate analyses, conducted with a variety of methodologies and looking at very different problems, into the common language of social evolution theory.

Acknowledgments

Literature Cited

Ajar, É. 2003. Analysis of disruptive selection in subdivided populations. BMC Evolutionary Biology 3:22.

Axelrod, R., R. A. Hammond, and A. Grafen. 2004. Altruism via kin-selection strategies that rely on arbitrary tags with which they coevolve. Evolution 58:1833–1838.

Barton, N. H. 1995. A general model for the evolution of recombination. Genetical Research 65:123–144.

———. 2000. Genetic hitchhiking. Philosophical Transactions of the Royal Society B: Biological Sciences 355:1553–1562.

Barton, N. H., and B. Charlesworth. 1998. Why sex and recombination? Science 281:1986–1990.

Barton, N. H., and K. S. Gale. 1993. Genetic analysis of hybrid zones. Pages 13–45 *in* R. G. Harrison, ed. Hybrid zones and the evolutionary process. Oxford University Press, Oxford.

Barton, N. H., and M. Turelli. 1987. Adaptive landscapes, genetic distance, and the evolution of quantitative characters. Genetical Research 49:157–174.

———. 1991. Natural and sexual selection on many loci. Genetics 127:229–255.

Billiard, S., and T. Lenormand. 2005. Evolution of migration under kin selection and local adaptation. Evolution 59:13–23.

Boyd, R., and P. J. Richerson. 1992. Punishment allows the evolution of cooperation (or anything else) in sizeable groups. Ethology and Sociobiology 13:171–195.

Boyd, R., H. Gintis, S. Bowles, and P. J. Richerson. 2003. The evolution of altruistic punishment. Proceedings of the National Academy of Sciences of the USA 100:3531–3535.

Brown, S. P. 1999. Co-operation and conflict in host-manipulating parasites. Proceedings of the Royal Society B: Biological Sciences 266:1899–1904.

Bulmer, M. G. 1994. Theoretical evolution and ecology. Sinauer, Sunderland, MA.

Charlesworth, B. 1978. Some models of the evolution of altruistic behaviour between siblings. Journal of Theoretical Biology 72:297–319.

———. 2000. Book review: *Levels of selection in evolution*. Heredity 84:493.

Christiansen, F. B. 1991. On conditions for evolutionary stability for a continuously varying character. American Naturalist 138:37–50.

———. 1999. Population genetics of multiple loci. Wiley, New York.

Crow, J. F., and M. Kimura. 1970. An introduction to population genetics theory. Harper & Row, New York.

Darwin, C. R. 1859. The origin of species. J. Murray, London.

Dawkins, R. 1976. The selfish gene. Oxford University Press, Oxford.

Dionisio, F., and I. Gordo. 2006. The tragedy of the commons, the public goods dilemma, and the meaning of rivalry and excludability in evolutionary biology. Evolutionary Ecology Research 8:321–332.

Dugatkin, L. A., and H. K. Reeve. 1994. Behavioural ecology and levels of selection: dissolving the group selection controversy. Advances in the Study of Behavior 23:101–133.

Edwards, A. W. F. 1994. The fundamental theorem of natural selection. Biological Reviews 69:443–474.

Eshel, I., and U. Motro. 1981. Kin selection and strong evolutionary stability of mutual help. Theoretical Population Biology 19:420–433.

Falconer, D. S. 1981. Introduction to quantitative genetics. 2nd ed. Longman, New York.

Fehr, E., and S. Gächter. 2000. Cooperation and punishment in public goods experiments. American Economic Review 90:980–994.

———. 2002. Altruistic punishment in humans. Nature 415:137–140.

Feldman, M. W., and L. L. Cavalli-Sforza. 1978. Darwinian selection and "altruism." Theoretical Population Biology 14:268–280.

Fisher, R. A. 1930. The genetical theory of natural selection. Clarendon, Oxford.

———. 1941. Average excess and average effect of a gene substitution. Annals of Eugenics 11:53–63.

Foster, K. R., and T. Wenseleers. 2006. A general model for the evolution of mutualisms. Journal of Evolutionary Biology 19:1283–1293.

Foster, K. R., T. Wenseleers, and F. L. W. Ratnieks. 2001. Spite: Hamilton's unproven theory. Annales Zoologici Fennici 38:229–238.

———. 2006. Kin selection is the key to altruism. Trends in Ecology & Evolution 21:57–60.

Frank, S. A. 1986a. The genetic value of sons and daughters. Heredity 56:351–354.

———. 1986b. Hierarchical selection theory and sex ratios. I. General solutions for structured populations. Theoretical Population Biology 29:312–342.

———. 1994. Genetics of mutualism: the evolution of altruism between species. Journal of Theoretical Biology 170:393–400.

———. 1995a. George Price's contributions to evolutionary genetics. Journal of Theoretical Biology 175:373–388.

———. 1995b. Mutual policing and repression of competition in the evolution of cooperative groups. Nature 377:520–522.

———. 1995c. The origin of synergistic symbiosis. Journal of Theoretical Biology 176:403–410.

———. 1996a. Models of parasite virulence. Quarterly Review of Biology 71:37–78.

———. 1996b. Policing and group cohesion when resources vary. Animal Behaviour 52:1163–1169.

———. 1997a. Multivariate analysis of correlated selection and kin selection, with an ESS maximization method. Journal of Theoretical Biology 189:307–316.

———. 1997b. 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.

———. 2003. Repression of competition and the evolution of cooperation. Evolution 57:693–705.

Gandon, S., and Y. Michalakis. 2000. Evolutionarily stable dispersal rate in a metapopulation with extinctions and kin competition. Journal of Theoretical Biology 199:275–290.

Gardner, A., and A. T. Kalinka. 2006. Recombination and the evolution of mutational robustness. Journal of Theoretical Biology 241:707–715.

Gardner, A., and S. A. West. 2004a. Cooperation and punishment, especially in humans. American Naturalist 164:753–764.

———. 2004b. Spite among siblings. Science 305:1413–1414.

———. 2004c. Spite and the scale of competition. Journal of Evolutionary Biology 17:1195–1203.

Gardner, A., S. A. West, and A. Buckling. 2004. Bacteriocins, spite and virulence. Proceedings of the Royal Society B: Biological Sciences 271:1529–1535.

Gardner, A., I. C. W. Hardy, P. D. Taylor, and S. A. West. 2007. Spiteful soldiers and sex ratio conflict in polyembryonic parasitoid wasps. American Naturalist (forthcoming).

Gintis, H. 2000. Strong reciprocity and human sociality. Journal of Theoretical Biology 206:169–179.

Grafen, A. 1985a. A geometric view of relatedness. Oxford Surveys in Evolutionary Biology 2:28–89.

———. 1985b. Hamilton's rule OK. Nature 318:310–311.

———. 2000. Developments of the Price equation and natural selection under uncertainty. Proceedings of the Royal Society B: Biological Sciences 267:1223–1227.

———. 2006a. Optimisation of inclusive fitness. Journal of Theoretical Biology 238:541–563.

———. 2006b. A theory of Fisher's reproductive value. Journal of Mathematical Biology 53:15–60.

Haldane, J. B. S. 1964. A defence of beanbag genetics. Perspectives in Biology and Medicine 19:343–359.

Hamilton, W. D. 1963. The evolution of altruistic behaviour. American Naturalist 97:354–356.

———. 1964. The genetical evolution of social behaviour. I. Journal of Theoretical Biology 7:1–16.

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

———. 1975. Innate social aptitudes of man: an approach from evolutionary genetics. Pages 133–153 in R. Fox, ed. Biosocial anthropology. Malaby, London.

———. 1996. Narrow roads of gene land. Vol. 1. Evolution of social behaviour. W. H. Freeman, Oxford.

Heinrich, J., and R. Boyd. 2001. Why people punish defectors. Journal of Theoretical Biology 208:79–89.

Herre, E. A. 1999. Laws governing species interactions? encouragement and caution from figs and their associates. Pages 209–237 in L. Keller, ed. Levels of selection in evolution. Princeton University Press, Princeton, NJ.

Hirshleifer, D., and E. Rasmusen. 1989. Cooperation in a repeated Prisoner's Dilemma with ostracism. Journal of Economic Behavior and Organization 12:87–106.

Hurst, L. D., A. Atlan, and B. O. Bengtsson. 1996. Genetic conflicts. Quarterly Review of Biology 71:317–364.

Johnstone, R. A. 2002. The evolution of inaccurate mimics. Nature 418:524–526.

Keller, L., and K. G. Ross. 1998 Selfish genes: a green beard in the red fire ant. Nature 394:573–575.

Kimura, M. 1965. Attainment of quasi linkage equilibrium when gene frequencies are changing by natural selection. Genetics 52:875–890.

Kirkpatrick, M., and M. J. Ryan. 1991. The evolution of mating preferences and the paradox of the lek. Nature 350:33–38.

Kirkpatrick, M., T. Johnson, and N. H. Barton. 2002. General models of multilocus evolution. Genetics 161:1727–1750.

Malécot, G. 1948. Les mathematiques de l'hérédité. Masson, Paris.

Maynard Smith, J. 1964. Group selection and kin selection. Nature 201:1145–1147.

———. 1978. The evolution of sex. Cambridge University Press, Cambridge.

Maynard Smith, J., and J. Haigh. 1974. Hitch-hiking effect of a favourable gene. Genetical Research 23:23–35.

Maynard Smith, J., and G. R. Price. 1973. The logic of animal conflict. Nature 246:15–18.

Maynard Smith, J., and E. Szathmáry. 1995. The major transitions in evolution. Oxford University Press, Oxford.

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

Nagylaki, T. 1993. The evolution of multilocus systems under weak selection. Genetics 134:627–647.

Nee, S. 1989. Does Hamilton's rule describe the evolution of reciprocal altruism? Journal of Theoretical Biology 141:81–91.

Otto, S. P., and M. W. Feldman. 1997. Deleterious mutations, variable epistatic interactions, and the evolution of recombination. Theoretical Population Biology 51:134–147.

Panchanathan, K., and R. Boyd. 2004. Indirect reciprocity can stabilize cooperation without the second-order free rider problem. Nature 432:499–502.

Pen, I., and F. J. Weissing. 2000. Towards a unified theory of cooperative breeding: the role of ecology and life history re-examined. Proceedings of the Royal Society B: Biological Sciences 267:2411–2418.

Pepper, J. W. 2000. Relatedness in trait group models of social evolution. Journal of Theoretical Biology 206:355–368.

Pepper, J. W., and B. B. Smuts. 2002. A mechanism for the evolution of altruism among nonkin: positive assortment through environmental feedback. American Naturalist 160:205–212.

Price, G. R. 1970. Selection and covariance. Nature 227:520–521.

———. 1972. Extension of covariance selection mathematics. Annals of Human Genetics 35:485–490.

———. 1995. The nature of selection. Journal of Theoretical Biology 175:389–396.

Price, G. R., and C. A. B. Smith. 1972. Fisher's Malthusian parameter and reproductive value. Annals of Human Genetics 36:1–7.

Provine, W. B. 1971. The origins of theoretical population genetics. University of Chicago Press, Chicago.

Queller, D. C. 1984. Kin selection and frequency dependence: a game-theoretic approach. Biological Journal of the Linnean Society 23:133–143.

———. 1985. Kinship, reciprocity, and synergism in the evolution of social behaviour. Nature 318:366–367.

———. 1992. A general model for kin selection. Evolution 46:376–380.

———. 2004. Kinship is relative. Nature 430:975.

Queller, D. C., E. Ponte, S. Bozzaro, and J. E. Strassmann. 2003. Single-gene greenbeard effects in the social amoeba *Dictyostelium discoideum*. Science 299:105–106.

Robertson, A. 1966. A mathematical model of the culling process in dairy cattle. Animal Production 8:95–108.

———. 1968. The spectrum of genetic variation. Pages 5–16 *in* R. C. Lewontin, ed. Population biology and evolution. Syracuse University Press, Syracuse, NY.

Rousset, F. 2004. Genetic structure and selection in subdivided populations. Princeton University Press, Princeton, NJ.

Roze, D., and F. Rousset. 2005. Inbreeding depression and the evolution of dispersal rates: a multilocus model. American Naturalist 166:708–721.

Schwartz, J. 2002. Population genetics and sociobiology: conflicting views of evolution. Perspectives in Biology and Medicine 45:224–240.

Sell, J., and R. K. Wilson. 1999. The maintenance of cooperation: expectations of future interaction and the trigger of group punishment. Social Forces 77:1551–1570.

Sober, E., and D. S. Wilson. 1998. Unto others: the evolution and psychology of unselfish behavior. Harvard University Press, Cambridge, MA.

Taylor, P. D. 1990. Allele-frequency change in a class-structured population. American Naturalist 135:95–106.

———. 1996. Inclusive fitness arguments in genetic models of behaviour. Journal of Mathematical Biology 34:654–674.

Taylor, P. D., and S. A. Frank. 1996. How to make a kin selection model. Journal of Theoretical Biology 180:27–37.

Taylor, P. D., G. Wild, and A. Gardner. 2007. Direct fitness or inclusive fitness: how shall we model kin selection? Journal of Evolutionary Biology 20:296–304. doi:1420-9101.2006.01196.x.

Trivers, R. L. 1985. Social evolution. Benjamin/Cummings, Menlo Park, CA.

Wade, M. J. 1985. Soft selection, hard selection, kin selection, and group selection. American Naturalist 125:61–73.

Wenseleers, T. 2001. Conflict from cell to colony. PhD diss. University of Leuven, Belgium.

———. 2006. Modelling social evolution: the relative merits and limitations of a Hamilton's rule–based approach. Journal of Evolutionary Biology 19:1419–1422.

Wenseleers, T., F. L. W. Ratnieks, and J. Billen. 2003. Caste fate conflict in swarm-founding social Hymenoptera: an inclusive fitness analysis. Journal of Evolutionary Biology 16:647–658.

Wenseleers, T., H. Helanterä, A. G. Hart, and F. L. W. Ratnieks. 2004. Worker reproduction and policing in insect societies: an ESS analysis. Journal of Evolutionary Biology 17:1035–1047.

West, S. A., and A. Buckling. 2003. Cooperation, virulence and siderophore production in bacterial parasites. Proceedings of the Royal Society B: Biological Sciences 270:37–44.

West, S. A., E. T. Kiers, I. Pen, and R. F. Denison. 2002. Sanctions and mutualism stability: when should less beneficial mutualists be tolerated? Journal of Evolutionary Biology 15:830–837.

West, S. A., A. S. Griffin, and A. Gardner. 2007. Social semantics: altruism, cooperation, mutualism and strong reciprocity. Journal of Evolutionary Biology. doi:1420-9101.2006.01258.x.

Appendix A from A. Gardner et al., "The Relation between Multilocus Population Genetics and Social Evolution Theory"

(Am. Nat., vol. 169, no. 2, p. 000)

Evolution of Cooperation

Here we derive results for the evolution of cooperation examples. We considered two models, described in the text and in figure 2. Note that the model described in figure $2B$ is a special case of that in figure $2C$, for when the synergy parameter is $d = 0$. Here we assume the more general model of figure $2C$ and reduce the results to the special case of figure $2B$ once we derive them. The generalized fitness function is given in expression (22). Substituting in allelic deviations ($X_i \rightarrow \zeta_i + p_i$), we have

$$w = 1 + (b - c + dp)p + (-c + dp)\zeta_1 + (b + dp)\zeta_2 + d\zeta_{12}. \tag{A1}$$

Following the usual procedure, we extract generalized selection coefficients:

$$\alpha_1 = \frac{-c + dp}{\bar{w}},$$

$$\alpha_2 = \frac{b + dp}{\bar{w}}, \tag{A2}$$

$$\alpha_{12} = \frac{d}{\bar{w}},$$

where mean fitness is

$$\bar{w} = 1 + (b - c)p + d(p^2 + \mathfrak{D}_{12}). \tag{A3}$$

Thus, the cooperation allele increases in frequency when

$$\Delta p = \sum_A \alpha_A \mathfrak{D}_{1A} = \alpha_1 pq + \alpha_2 \mathfrak{D}_{12} + \alpha_{12}(1 - 2p)\mathfrak{D}_{12} > 0. \tag{A4}$$

Dividing by pq and noticing that $\mathfrak{D}_{12}/pq = \beta_{X_2, X_1}$ is the kin selection coefficient of relatedness, we have the condition

$$\alpha_1 + \alpha_2 \beta_{X_2, X_1} + \alpha_{12}(1 - 2p)\beta_{X_2, X_1} > 0. \tag{A5}$$

Substituting the model parameters gives

$$\frac{-c + dp}{\bar{w}} + \frac{b + dp}{\bar{w}}R + \frac{d}{\bar{w}}(1 - 2p)R > 0, \tag{A6}$$

or

$$R(b + dp) - c + dp + (1 - 2p)Rd > 0. \tag{A7}$$

Note that, for the special case when $d = 0$, we have

1

$$Rb - c > 0. \tag{A8}$$

Inequality (A7) may be rearranged into the usual Hamilton's rule form, $RB - C > 0$, where $-C = \beta_{w,x_1 \cdot x_2}$, $B = \beta_{w,x_2 \cdot x_1}$, and $R = \beta_{x_2, x_1}$, in accordance with the canonical derivation of Hamilton's rule from the STNS (eqq. [13], [14]). Here, the cost and benefit are found in the standard way, by determining the partial-regression terms that minimize the residual variance that is not explained by fitting the additive linear model of expected fitness as a function of genic values (X_1 and X_2),

$$w_{\exp} = \bar{w} + \beta_{w,X_1 \cdot X_2}(X_1 - p) + \beta_{w,X_2 \cdot X_1}(X_2 - p), \tag{A9}$$

to the population. The unexplained variance is given by $\mathrm{E}\big[(w_{\exp} - w_{\mathrm{obs}})^2\big]$, where w_{obs} is the observed, or actual, fitness. In this model, individual fitness can take one of four possible values, depending on the genotype of the social pairing: (1) with probability $q^2 + Rpq$, both individuals are defectors, in which case, $X_1 = X_2 = 0$, and while the fitness of the focal individual is $w_{\mathrm{obs}} = 1$, the fitness predicted by the model in equation (A9) is $w_{\exp} = 1 - (\beta_{w,x_1 \cdot x_2} + \beta_{w,x_2 \cdot x_1})p$; (2) with probability $pq(1 - R)$, the focal individual is a cooperator and its partner is a defector, in which case, $X_1 = 1$, $X_2 = 0$, fitness of the focal individual is $w_{\mathrm{obs}} = 1 - c$, and the predicted fitness is $w_{\exp} = 1 + \beta_{w,x_1 \cdot x_2}(1 - p) - \beta_{w,x_2 \cdot x_1}p$; (3) with probability $pq(1 - R)$, the focal individual is a defector and its partner is a cooperator, in which case, $X_1 = 0$, $X_2 = 1$, fitness of the focal individual is $w_{\mathrm{obs}} = 1 + b$, and the predicted fitness is $w_{\exp} = 1 - \beta_{w,x_1 \cdot x_2}p + \beta_{w,x_2 \cdot x_1}(1 - p)$; (4) with probability $p^2 + Rpq$, both individuals are cooperators, in which case, $X_1 = X_2 = 1$, the fitness of the focal individual is $w_{\mathrm{obs}} = 1 - c + b + d$, and the predicted fitness is $w_{\exp} = 1 + (\beta_{w,x_1 \cdot x_2} + \beta_{w,x_2 \cdot x_1})(1 - p)$. Using this, it is straightforward to show that the unexplained variance is minimized when the model parameters are given by

$$\beta_{w,X_1 \cdot X_2} = -c + \frac{R + (1 - R)p}{1 + R}d,$$

$$\beta_{w,X_2 \cdot X_1} = b + \frac{R + (1 - R)p}{1 + R}d, \tag{A10}$$

which implies that $B = b + \{[R + (1 - R)p]/(1 + R)\}d$ and $C = c - \{[R + (1 - R)p]/(1 + R)\}d$. Having obtained these fitness components, it easy to show that the expression $RB - C$ recovers the left-hand side of inequality (A7).

Appendix B from A. Gardner et al., "The Relation between Multilocus Population Genetics and Social Evolution Theory"
(Am. Nat., vol. 169, no. 2, p. 000)

Cooperation and Punishment

Here we present the analysis for the cooperation and punishment example. In the text, we proposed a fitness function, equation (25). Making the substitution $X_i \to \zeta_i + p_i$ gives

$$w = 1 + \frac{b-c}{2}p_i - \frac{(a+d)q_i + e}{2}p_j + \frac{dp_j - c}{2}\zeta_{i_1} + \frac{ap_j + b}{2}\zeta_{i_2} - \frac{aq_i + e}{2}\zeta_{j_1} - \frac{dq_i}{2}\zeta_{j_2} + \frac{d}{2}\zeta_{i_1 j_2} + \frac{a}{2}\zeta_{i_2 j_1}. \quad (B1)$$

From this, we can extract the multilocus selection coefficients:

$$\alpha_{i_1} = \frac{dp_j - c}{2\bar{w}},$$

$$\alpha_{i_2} = \frac{ap_j + b}{2\bar{w}},$$

$$\alpha_{j_1} = -\frac{aq_i + e}{2\bar{w}},$$

$$\alpha_{i_2} = -\frac{dq_i}{2\bar{w}}, \quad (B2)$$

$$\alpha_{i_1 j_2} = \frac{d}{2\bar{w}},$$

$$\alpha_{i_2 j_1} = \frac{a}{2\bar{w}}.$$

The population mean fitness ($\bar{w}$) is found by taking the expectation of fitness (w),

$$\bar{w} = 1 + \frac{b-c}{2}p_i - \frac{(a+d)q_i + e}{2}p_j + R\frac{a+d}{2}\mathfrak{D}_{ij}, \quad (B3)$$

where $\mathfrak{D}_{i_1 j_2} = \mathfrak{D}_{i_2 j_1} = R\mathfrak{D}_{ij}$ and $\mathfrak{D}_{ij} = \mathfrak{D}_{i_1 j_1} = \mathfrak{D}_{i_2 j_2}$ is the association between loci within individuals. Selection operating on the cooperation locus changes the allele frequency according to the recursion

$$p_i' = p_i + \sum_{\mathbf{A}} \alpha_A \mathfrak{D}_{Ai_1} = (\alpha_{i_1} + R\alpha_{i_2})p_i q_i + (\alpha_{j_1} + R\alpha_{j_2})\mathfrak{D}_{ij} + R(\alpha_{i_1 j_2} + \alpha_{i_2 j_1})(1 - 2p_i)\mathfrak{D}_{ij}, \quad (B4)$$

and at the punishment locus, allele frequency change is given by

$$p_j' = p_j + \sum_{\mathbf{A}} \alpha_A \mathfrak{D}_{Aj_1} = (\alpha_{j_1} + R\alpha_{j_2})p_j q_j + (\alpha_{i_1} + R\alpha_{i_2})\mathfrak{D}_{ij} + R(\alpha_{i_1 j_2} + \alpha_{i_2 j_1})(1 - 2p_j)\mathfrak{D}_{ij}. \quad (B5)$$

The association within individuals, between loci, is shaped by selection and recombination and is given by the recursion

1

$$\mathfrak{D}'_{ij} = (1 - r)\Big[\mathfrak{D}_{ij} + \sum_A \mathfrak{a}_A(\mathfrak{D}_{Ai_1j_1} - \mathfrak{D}_A\mathfrak{D}_{ij}) - (p'_i - p_i)(p'_j - p_j)\Big]$$

$$= (1 - r)\{\mathfrak{D}_{ij} + (\mathfrak{a}_{i_1} + R\mathfrak{a}_{i_2})(1 - 2p_i)\mathfrak{D}_{ij} + (\mathfrak{a}_{j_1} + R\mathfrak{a}_{j_2})(1 - 2p_j)\mathfrak{D}_{ij} \tag{B6}$$

$$+ R(\mathfrak{a}_{i_1j_2} + \mathfrak{a}_{i_2j_1})[p_iq_ip_jq_j + (1 - 2p_1)(1 - 2p_j)\mathfrak{D}_{ij} - \mathfrak{D}^2_{ij}] - (p'_i - p_i)(p'_j - p_j)\}.$$

We now ask, under which circumstances is a population monomorphic for defection and forgiveness resistant to invasion by cooperation and punishment? We can linearize our recursions in p_i, p_j, and $\mathfrak{D}_{ij}$ by having p_i and p_j take vanishingly small values, and hence $\mathfrak{D}_{ij}$ will also be vanishingly small, and so we may ignore higher-order terms. This gives us recursions of the form

$$\begin{pmatrix} p'_i \\ p'_j \\ \mathfrak{D}'_{ij} \end{pmatrix} = \begin{pmatrix} \alpha_1 & 0 & \alpha_3 \\ 0 & \alpha_5 & \alpha_6 \\ 0 & 0 & \alpha_9 \end{pmatrix} \begin{pmatrix} p_i \\ p_j \\ \mathfrak{D}_{ij} \end{pmatrix}, \tag{B7}$$

and so the eigenvalues (λ) are given by the diagonal coefficients

$$\lambda_1 = \alpha_1 = 1 + \frac{Rb - c}{2},$$

$$\lambda_2 = \alpha_5 = 1 - \frac{a + Rd + e}{2}, \tag{B8}$$

$$\lambda_3 = \alpha_9 = (1 - r)\Big[1 - \frac{(1 - R)(a + c + d) + e}{2}\Big].$$

From λ_1, in the absence of the punishment allele, the condition for cooperation to invade is $Rb - c > 0$. Since we are concerned only with situations where punishment is required in order for cooperation to be favored, we proceed on the assumption that $Rb - c < 0$. Hence, all three eigenvalues have magnitude less than unity, and so perturbations to the monomorphic population are neutralized. Therefore, cooperation and punishment cannot invade the population.

We now consider whether a population monomorphic for cooperation and punishment is resistant to invasion by defection and forgiveness. We repeat the same procedure as before, although now we follow the frequencies of the vanishingly rare defection and forgiveness alleles, and determine the three eigenvalues

$$\lambda_1 = \alpha_1 = 1 - \frac{R(a + b) + d - c}{2 + b - c - e},$$

$$\lambda_2 = \alpha_5 = 1 + \frac{e}{2 + b - c - e}, \tag{B9}$$

$$\lambda_3 = \alpha_9 = (1 - r)\Big[1 - \frac{Rb - c + (1 - R)d - e}{2 + b - c - e}\Big].$$

All three eigenvalues are positive. If there is any intrinsic cost of punishing ($e > 0$), then $\lambda_2 > 1$, and hence the leading eigenvalue exceeds unity and the cooperating-punishing population is invadable. If there is no intrinsic cost of punishing ($e = 0$), then a sufficient condition for invasion (from λ_1) is $R(a + b) + d < c$, because when this is satisfied, cooperation is not favored even when all individuals punish. Assuming that this is not satisfied, that is, that punishment is a deterrent, then cooperation-punishment is neutrally stable when $\lambda_3 < 1$, which is true for sufficiently frequent recombination (large enough r). Under no circumstances does selection maintain punishment.

Appendix C from A. Gardner et al., "The Relation between Multilocus Population Genetics and Social Evolution Theory"

(Am. Nat., vol. 169, no. 2, p. 000)

Evolution of Relatedness

Here we present the analysis for the evolution of relatedness example. In the model outlined in the text, the genetic structure of two paired populations is changed, first, by reproduction and, second, by dispersal. Taking a levels-of-selection approach, we model reproduction as two separate group selection and individual selection (genotype-dependent transmission) steps, and these are followed by the dispersal (genotype-independent transmission) step. We begin with the group fitness function (26), derived in the text. Substituting in allelic values, we have

$$w = 1 + 2sp + (s^2 + \varepsilon)p^2 + [s + (s^2 + \varepsilon)p](\zeta_1 + \zeta_2) + (s^2 + \varepsilon)\zeta_{12}, \tag{C1}$$

and hence we are able to extract generalized selection coefficients

$$\alpha_1 = \alpha_2 = \frac{s + (s^2 + \varepsilon)p}{\bar{w}},$$

$$\alpha_{12} = \frac{s^2 + \varepsilon}{\bar{w}}, \tag{C2}$$

where population mean fitness is given by

$$\bar{w} = 1 + 2sp + (s^2 + \varepsilon)(p^2 + \mathfrak{D}_{12}). \tag{C3}$$

The change in the cooperation allele during the group selection phase is described by

$$p' = p + \sum_A \alpha_A \mathfrak{D}_{A1} = p + \alpha_1 pq + \alpha_2 \mathfrak{D}_{12} + \alpha_{12}(1 - 2p)\mathfrak{D}_{12}, \tag{C4}$$

and, correcting for the change in reference value, the association between individuals after group selection is

$$\mathfrak{D}'_{12} = \mathfrak{D}_{12} + \sum_A \alpha_A(\mathfrak{D}_{A12} - \mathfrak{D}_A \mathfrak{D}_{12}) - (p' - p)^2$$

$$= \mathfrak{D}_{12} + (\alpha_1 + \alpha_2)(1 - 2p)\mathfrak{D}_{12} + \alpha_{12}[p^2 q^2 + (1 - 2p)^2 \mathfrak{D}_{12} - \mathfrak{D}_{12}^2] - (p' - p)^2. \tag{C5}$$

Note that here we are describing the genetic composition of the paired populations after group selection but before transmission and that hence there are a number of cooperator individuals who will effectively be transformed into defectors during the transmission process.

We now describe the change due to genotype-dependent transmission. Previously, we derived transmission functions for the genotype-dependent transmission event (eq. [27]). For each of these functions, we may substitute allelic deviation terms and rearrange so as to recover the form of expression (7), giving us generalized transmission coefficients. From the first function, we have

$$\bar{t}_{i \leftarrow i} = 1 - \frac{f}{1+s}(p'q' - \mathcal{D}'_{12}),$$

$$\tau_{i \leftarrow i | i} = -\frac{f}{1+s}q',$$

$$\tau_{i \leftarrow i | j} = \frac{f}{1+s}p', \tag{C6}$$

$$\tau_{i \leftarrow i | \{i,j\}} = \frac{f}{1+s}.$$

The first expression is the average probability that the species-*i* individual in the daughter group derives from the species-*i* individual in the parent group; the second expression is the effect on this probability, given that the species-*i* individual in the parent group is a cooperator; the third expression is the effect on the probability, given that the species-*j* individual is a cooperator; and the fourth expression is the effect on the probability, given that both individuals in the parent group are cooperators. From the second function, we have

$$\bar{t}_{i \leftarrow k} = \frac{f}{1+s}(p'q' - \mathcal{D}'_{12}),$$

$$\tau_{i \leftarrow k | i} = \frac{f}{1+s}q',$$

$$\tau_{i \leftarrow k | j} = -\frac{f}{1+s}p', \tag{C7}$$

$$\tau_{i \leftarrow k | \{i,j\}} = -\frac{f}{1+s};$$

from the third function, we have

$$\bar{t}_{\{i,j\} \leftarrow \{i,j\}} = 1 - 2\frac{f}{1+s}(p'q' - \mathcal{D}'_{12}),$$

$$\tau_{\{i,j\} \leftarrow \{i,j\} | i} = -(1 - 2p')\frac{f}{1+s},$$

$$\tau_{\{i,j\} \leftarrow \{i,j\} | j} = -(1 - 2p')\frac{f}{1+s}p', \tag{C8}$$

$$\tau_{\{i,j\} \leftarrow \{i,j\} | \{i,j\}} = 2\frac{f}{1+s};$$

and from the fourth function, we have

$$\bar{t}_{\{i,j\} \leftarrow \{k,j\}} = \frac{f}{1+s}(p'q' - \mathcal{D}'_{12}),$$

$$\tau_{\{i,j\} \leftarrow \{k,j\} | i} = \frac{f}{1+s}q',$$

$$\tau_{\{i,j\} \leftarrow \{k,j\} | j} = -\frac{f}{1+s}p', \tag{C9}$$

$$\tau_{\{i,j\} \leftarrow \{k,j\} | \{i,j\}} = -\frac{f}{1+s}.$$

2

Substituting into expression (8) and noting that, for example, $\tau_{1\leftarrow k|2} = \tau_{2\leftarrow k|1} = \tau_{i\leftarrow k|j}$, we find that the change in the allele frequency due to genotype-dependent transmission is given by

$$p'' = p' - \frac{f}{1+s}(p'q' - \mathcal{D}'_{12}), \tag{C10}$$

and the change in the association between social partners due to genotype-dependent transmission, after the reference values have been updated, is

$$\mathcal{D}''_{ij} = \mathcal{D}'_{ij} + p'^2 - p''^2. \tag{C11}$$

The genotype-independent transmission that occurs because of random dispersal does not alter allele frequencies ($p''' = p''$), but it does reduce between-individual associations in the usual way:

$$\mathcal{D}'''_{12} = (1-m)\mathcal{D}''_{12}. \tag{C12}$$

Thus, we have formulated exact recursions describing the change in allele frequency and association between social partners over the course of one generation. We now proceed to an invasion analysis, to ask how the system evolves, given an initially vanishing frequency of cooperators. We linearize the recursions for allele frequency and association between social partners because higher-order terms become negligible when these are vanishingly small, giving

$$\begin{pmatrix} p''' \\ \mathcal{D}'''_{12} \end{pmatrix} = \begin{pmatrix} \alpha_1 & \alpha_2 \\ 0 & \alpha_4 \end{pmatrix} \begin{pmatrix} p \\ \mathcal{D}_{12} \end{pmatrix}, \tag{C13}$$

and hence the eigenvalues of this system are given by

$$\lambda_1 = \alpha_1 = 1 + s - f,$$
$$\lambda_2 = \alpha_4 = (1-m)[(1+s)^2 + \varepsilon]. \tag{C14}$$

If $1 + s - f > 1$, then at least one of the eigenvalues (λ_1) is greater than unity, and hence cooperation invades when rare. If $1 + s - f < 1$, then the cooperation allele is disfavored when $\lambda_2 < 1$ and favored when $\lambda_2 > 1$. When λ_2 is leading, the state of the system during the asymptotic ascent of the cooperation allele is given by the right eigenvector

$$\begin{pmatrix} p \\ \mathcal{D}_{12} \end{pmatrix} = \left(\begin{matrix} p \\ \left[1 - m\left(1 + \dfrac{1+s-f}{s+s^2+\varepsilon+f} \right) \right]p \end{matrix} \right). \tag{C15}$$

On the assumption that λ_2 is leading, we may interpret the asymptotic rate of increase in terms of Hamilton's rule, that is, $\lambda_2 = 1 + RB - C$, where the components of the rule have their usual definitions. This Hamilton's rule is therefore given as $(1-m)[(1+s)^2 + \varepsilon] > 1$. The relatedness coefficient may be determined from the usual regression definition, that is, $\beta_{X_2, X_1} = \mathcal{D}_{12}/pq$, and from the dominant eigenvector (C15), this gives $R \to 1 - m\{1 + [(1+s-f)/(s+s^2+\varepsilon+f)]\}$ as $p \to 0$. The cost and benefit terms are less easy to discern in this levels-of-selection approach, but they are readily recovered by writing a neighbor-modulated fitness for the individuals in the model,

$$w = X_1 X_2[(1+s)^2 + \varepsilon] + X_1(1-X_2)(1+s-f) + (1-X_1)X_2(1+s+f) + (1-X_1)(1-X_2), \tag{C16}$$

and applying the linear-model-fitting approach outlined in appendix A. The partial-regression coefficients minimizing the unexplained variance are, in the limit of $p \to 0$,

$$\beta_{w,X_1 \cdot X_2} = s - f + \frac{R}{1+R}(s^2 + \varepsilon) = -C,$$

$$\beta_{w,X_2 \cdot X_1} = s + f + \frac{R}{1+R}(s^2 + \varepsilon) = B, \tag{C17}$$

and substituting these into the asymptotic rate of increase $1 + RB - C$ recovers the leading eigenvalue λ_2.