Biol Philos (2014) 29:477–515
DOI 10.1007/s10539-013-9404-0

Inclusive fitness and the sociobiology of the genome

Herbert Gintis

Received: 23 January 2013 / Accepted: 9 November 2013 / Published online: 1 December 2013

Abstract Inclusive fitness theory provides conditions for the evolutionary success of a gene. These conditions ensure that the gene is *selfish* in the sense of Dawkins (The selfish gene, Oxford University Press, Oxford, 1976): genes do not and cannot sacrifice their own fitness on behalf of the reproductive population. Therefore, while natural selection explains the appearance of design in the living world (Dawkins in The blind watchmaker: why the evidence of evolution reveals a universe without design, W. W. Norton, New York, 1996), inclusive fitness theory does not explain how. Indeed, Hamilton's rule is equally compatible with the evolutionary success of prosocial altruistic genes and antisocial predatory genes, whereas only the former, which account for the appearance of design, predominate in successful organisms. Inclusive fitness theory, however, permits a formulation of the central problem of sociobiology in a particularly poignant form: how do interactions among loci induce utterly selfish genes to collaborate, or to predispose their carriers to collaborate, in promoting the fitness of their carriers? Inclusive fitness theory, because it abstracts from synergistic interactions among loci, does not answer this question. Fitness-enhancing collaboration among loci in the genome of a reproductive population requires suppressing alleles that decrease, and promoting alleles that increase the fitness of its carriers. Suppression and promotion are effected by regulatory networks of genes, each of which is itself utterly selfish. This implies that genes, and a fortiori individuals in a social species, do not maximize inclusive fitness but rather interact strategically in complex ways. It is the task of sociobiology to model these complex interactions.

Keywords Inclusive fitness · Hamilton's rule · Sociobiology

H. Gintis (✉)
Santa Fe Institute, 1399 Hyde Park Rd, Santa Fe, NM 87501, USA
e-mail: hgintis@comcast.net

> One can recognize in the evolution of life several revolutions... In each of these revolutions, there has been a conflict between selection at several levels. The achievement of individuality at the higher level has required that the disruptive effects of selection at the lower level be suppressed.
>
> Maynard Smith (1988)

> A Mendelian population has a common gene pool, which is its collective or corporate genotype.
>
> Dobzhansky (1953)

Introduction

Hamilton's rule provides conditions for the evolutionary success of a gene. These conditions ensure that the gene is *selfish* in the sense described by Dawkins (1976). In particular, Hamilton's rule implies that the conditions for the evolutionary success of a gene are distinct from the conditions under which the gene enhances the mean fitness of the reproductive population. Therefore inclusive fitness does not explain the appearance of design in nature or, in other words, why the genome of a successful species consists of genes that predominantly collaborate in promoting the fitness of its members (Dawkins 1996). Indeed, Hamilton's rule equally supports the evolutionary success of prosocial altruistic genes and antisocial predatory genes (these terms are formally defined below), whereas the former predominate in a successful species and account for the appearance of design.

It is common for sociobiologists who, with Dawkins, adopt the "gene's-eye point of view" to overlook or deny this fact, despite its being a simple logical implication of Hamilton's rule. Indeed, it many population biologist claim that the appearance of design in nature is explained by Hamilton's rule. For instance, a letter to the journal *Nature*, signed by 137 prominent biological researchers protesting a critique of inclusive fitness theory by Nowak et al. (2010), asserts:

> Natural selection explains the appearance of design in the living world, and inclusive fitness theory explains what this design is for. Specifically, natural selection leads organisms to become adapted as if to maximize their inclusive fitness (Abbot et al. 2011).

In fact, organisms do not generally maximize inclusive fitness. Rather, organisms in a social species interact strategically in a complex manner involving collaboration, as well as promotion and suppression of behaviors. Moreover, relatedness may play a subordinate role in the dynamics of a species, especially one that exhibits a complex division of labor.

Hamilton's rule

Classical genetics does not model cases in which individuals sacrifice on behalf of non-offspring, such as sterile workers in an insect colony (Wheeler 1928),

cooperative breeding in birds (Skutch 1961), and altruistic behavior in humans (Darwin 1871). This problem was addressed by Hamilton (1963, 1964a, b, 1970), who noted that if a gene favorable to helping others is likely to be present in the recipient of an altruistic act, then the gene could evolve even if it reduces the fitness of the donor. Hamilton called this *inclusive fitness* theory.

Hamilton developed a simple inequality, operating at the level of a single locus, that gives the conditions for the evolutionary success of an allele. Thus rule says that if an allele in individual A increases the fitness of individual B whose degree of relatedness to A is r, and if the cost to A is c, while the fitness benefit to B is b, then the allele will evolve (grow in frequency in the reproductive population) if

$$br > c. \tag{1}$$

We call $br - c$ the *inclusive fitness* of the focal allele. Subsequent research supported some of Hamilton's major predictions (Maynard Smith and Ridpath 1972; Brown 1974; West-Eberard 1975; Krackauer 2005).

A critical appreciation of Hamilton's rule requires understanding when and why it is true. The rigorous derivations of Hamilton's rule (Hamilton 1964a; Grafen 1985; Queller 1992; Frank 1998) are mathematically sophisticated and difficult to interpret. The usual popular argument (for instance, Bourke 2011) assumes that an altruistic helping behavior (b, $c > 0$) is governed by an allele at a single locus, and r is the probability that the recipient of the help has a copy of the focal allele. The net fitness increment to carriers of the focal allele is then $br - c$, so the allele increases in frequency if this expression is positive.[1]

This derivation has the attractive side effect of showing that an evolutionarily successful allele is necessarily prosocial: by helping copies of itself, an altruistic allele helps the reproductive population as a whole. Indeed, the net change in population fitness from the helping behavior is $b - c$, which is then necessarily positive, since

$$b - c = b(1 - r) + (br - c) > b(1 - r) \geq 0. \tag{2}$$

However attractive, the popular argument has key weaknesses that render it, and some of the conclusions drawn from it, unacceptable.

- The argument shows the conditions for an increase in the *absolute number* of copies of the focal allele in the population, but says nothing about its *relative frequency*, which is the quantity relevant to the evolutionary success of the focal allele. Indeed, $br - c$, the net increase in the number of copies of the focal allele, is less than $b - c$, which is the net increase in the number of copies of all alleles at the focal locus, so the frequency of the focal allele in the next period will be lower than $br - c$, and may even decrease.
- Hamilton's rule can be satisfied with b, $c < 0$, in which case the focal allele can increase its frequency in the population by helping itself ($c < 0$) while harming

[1] Hamilton's rule extends directly to behavior that is governed by alleles at multiple loci, provided that the interactions among the loci are frequency independent, or equivalently, that the effects at distinct loci contribute additively to the phenotypic behavior. Grafen (1984) calls a such a phenotype a *p-score*. In this paper I will use the term "single locus" even in places where the *p*-score generalization applies.

others ($b < 0$). Such an allele may also be *antisocial*, imposing net harm on the reproductive population. For instance, suppose $b = -0.4$, $c = -0.3$ and $r = 1/2$. Then $br - c = 0.1 > 0$, so Hamilton's rule is satisfied, but $b - c = -0.1$, so the allele is antisocial. In fact, the focal allele in this case is evolutionarily successful yet antisocial for any $r < 3/4$.

- The derivation does not explain the relationship between r as a probability and r as a coefficient of relatedness. In fact, there is no simple relationship between the r in Hamilton's rule and genealogical coefficients of relatedness. The appropriate value of r in Hamilton's rule is generally a complex function of the social structure of the species in question.

To address these deficiencies, we begin our study of inclusive fitness theory with a careful derivation of Hamilton's rule in the case of a haploid species. A very general derivation in the diploid case is presented in Appendix 5.

Suppose there is an allele at a locus of the genome of a reproductive population that induces carrier A (called the *donor*) to incur a fitness change c that leads to a fitness change b in individual B (called the *recipient*). We will represent B as an individual, but in fact, the fitness change b can be spread over any number of individuals. If $b > 0$, A bestows a gain upon B, and if $c > 0$, A experiences a fitness loss. However, in general we make no presumption concerning the signs or magnitudes of b and c, except that selection is weak in the sense that b and c do not change, and the population does not become extinct, over the course of a single reproduction period. This assumption, which is extremely plausible, will be made throughout this paper.

Suppose the frequency of the focal allele in the population is q, where $0 < q < 1$, and the probability that B has a copy of the allele is p. Then if the size of the population is n, there are qn individuals with the focal allele, they change the number of members of the population from n to $n + qn(b - c)$, and they change the number of focal alleles from qn to $qn + qn(pb - c)$. Thus the frequency of the allele from one period to the next will increase if

$$\Delta q = \frac{qn + qn(pb - c)}{n + qn(b - c)} - q = \frac{q(1 - q)}{1 + q(b - c)}\left(b\frac{p - q}{1 - q} - c\right) > 0. \tag{3}$$

The condition for an increase in the focal allele thus is

$$b\left(\frac{p - q}{1 - q}\right) > c. \tag{4}$$

By contrast, the focal allele is *prosocial*, meaning that contributes to the fitness of the reproductive population, when $n + qn(b - c) > n$, or more simply $b > c$.

The equivalence of (4) with Hamilton's rule (1) requires

$$p = r + (1 - r)q. \tag{5}$$

In this equation, r is commonly interpreted as the probability that an agent A with focal allele a will encounter another agent B who also has the allele, through some social regularity specific to the species. With probability $1 - r$, it is assumed that the recipient B is a random member of the population, so B has the focal allele with a probability equal to the frequency q of the allele in the population.

Often (5) is explained using the concept of identity by descent (Malécot 1948; Crow 1954), where r is the probability that both donor A and recipient B have inherited the same focal allele from a common ancestor. For instance, if A and B are full siblings, then $r = 1/2$ because this is the probability that both have inherited the focal allele from the same parent. Moreover, if the siblings have inherited the focal allele from different parents, then they will still be the same allele with a probability equal to the frequency of the allele in the population, assuming no assortative mating. In general, r will then be the expected degree of identity by descent of recipients.

The rather stunning conclusion from this exercise in elementary algebra and gene-counting is what we may call the Inclusive Fitness Harmony Principle. To state this principle succinctly, we say the allele a is *helpful* if its carriers enhance the fitness of other individuals that it encounters ($b > 0$), *altruistic* if it incurs a fitness cost in so doing ($c > 0$), and *prosocial* if increases mean population fitness. We then have:

Inclusive Fitness Harmony Principle An evolutionarily successful ($br > c$) gene that is helpful ($b > 0$), and associates preferentially with relatives ($r > 0$) is necessarily prosocial ($b - c > 0$).

Note that when carriers interact only with other carriers (so-called *greenbeards*), then $r = 1$, so Hamilton's rule (1) is satisfied exactly when $b > c$. The fact that greenbeards are necessarily prosocial in the sense of increasing mean population fitness was first pointed out by Ridley and Grafen (1981). By contrast, if interaction is random, so $p = q$ in (4), or $r = 0$ in (1), then Hamilton's rule can never be satisfied when the cost c is positive, but always satisfied if c is negative, in which case the allele is again necessarily prosocial.

The importance of this principle for sociobiology is inestimable, and mirrors similar assertions concerning the social value of selfishness in humans offered by Bernard Mandeville in his famous *Fable of the Bees* (1705), in which "private vices" give rise to "public virtues," and Adam Smith's (1776) equally famous dictum, "It is not from the benevolence of the butcher, the brewer, or the baker that we expect our dinner, but from their regard to their own interest." While economists have determined the precise conditions—they are far from universal—under which Mandeville and Smith are correct (Mas-Colell et al. 1995), the Inclusive Fitness Harmony Principle is true under much broader conditions.

However, the assumption in the Inclusive Fitness Harmony Principle that the gene is helpful leaves open the question as to the evolutionary status of genes that are *not* helpful. An alternative to Principle is, indeed, prima facie equally possible. Suppose the focal allele is *predatory* in the sense that its carriers benefit ($c < 0$), and it imposes costs on other individuals ($b < 0$). Then Hamilton's rule becomes $(-b)r < (-c)$, which can be satisfied even though the focal allele is *antisocial* in the sense that $b - c < 0$. Indeed, this will be the case whenever $|b|(1 - r) > c - b > 0$. We conclude from the above reasoning:

Inclusive Fitness Disharmony Principle A gene that is evolutionarily successful ($br > c$), predatory ($b,c < 0$), and associates preferentially with relatives ($r > 0$) may be antisocial ($b - c < 0$).

Moreover, in our basic Eq. (4) for Hamilton's rule, $p < q$ is prima facie as likely as $p > q$. In the case $p < q$, we see that r in (5) must be negative, so r cannot be interpreted as a probability at all. In particular, r cannot then be interpreted as a degree of relatedness in the genealogical sense. Indeed, it is a common occurrence that the interaction is costly but involves reducing the fitness of others, and (5) can hold with $r < 0$, while the focal allele is still altruistic (Bourke 2011). Examples are warfare in ants (Hölldobler and Wilson 1990) and humans (Bowles and Gintis 2011), as well as generally spiteful behavior in many species (Hamilton 1970; Foster et al. 2001; Gardner et al. 2004).

Hamilton's rule can be extended to deal with heterogeneous relatedness, diploidy, dominance, coordinated cooperation, local resource competition, inbreeding, and other complications (Uyenoyama and Feldman 1980; Michod and Hamilton 1980; Queller 1992; Wilson et al. 1992; Taylor 1992; Rousset and Billard 2007), with an equation closely resembling (1) continuing to hold. In general, however, the frequency q will appear in (1), and b and c may be functions of q as well, so the interpretation of r as relatedness becomes accordingly more complex (Michod and Hamilton 1980).

Of course, in general b and c will also depend on the frequency of alleles at other loci of the genome, and since the change in frequency q of the focal allele in the population will affect the relative fitnesses of alleles at other loci, inducing changes in frequency at these loci, which in turn will affect the values of b, c, and even r. For this reason, Hamilton's rule presupposes *weak selection*, in the sense that population gene frequencies do not change appreciably in a single reproduction period. Therefore Hamilton's rule does not imply that a successful allele will move to fixation in the genome. Moreover, alleles at other loci that are enhanced in inclusive fitness by the focal allele's expansion may undergo mutations that enhance the inclusive fitness of the focal allele, while alleles at other loci that are harmed by the expansion of the focal allele may develop mutations that suppress the focal allele.

Inclusive fitness and kin selection

William Hamilton's early work in inclusive fitness focused on the role of genealogical kinship in promoting prosocial behavior. Hamilton writes, in his first full presentation of inclusive fitness theory (Hamilton 1964a, p. 19):

> In the hope that it may provide a useful summary, we therefore hazard the following generalized unrigorous statement of the main principle that has emerged from the model. The social behaviour of a species evolves in such a way that in each distinct behaviour-evoking situation the individual will seem to value his neighbours' fitness against his own according to the coefficients of relationship appropriate to that situation.

Because of this close association between inclusive fitness and the social relations among genealogical relatives, Maynard Smith (1964) called Hamilton's theory *kin selection*, by which he meant that individuals are predisposed to sacrifice on behalf of highly related family members.

A decade after Hamilton's seminal inclusive fitness papers, motivated by new empirical evidence and Price's equation (Price 1970), Hamilton (1975, p. 337) revised his views, writing:

> Kinship should be considered just one way of getting positive regression of genotype... the inclusive fitness concept is more general than 'kin selection'.

Nevertheless the two concepts are often equated, even in the technical literature. For instance, throughout his authoritative presentation of sex allocation theory, West (2009) identifies inclusive fitness with kin selection in several places and never distinguishes between the two terms at any point in the book. Similarly, in Bourke's (2011) ambitious introduction to sociobiology, we find:

> The basic theory underpinning social evolution [is] Hamilton's inclusive fitness theory (kin selection theory).

This curious identification of inclusive fitness theory, which models the dynamics at a single genetic locus and is equally at home with altruistic and predatory genes, with kin selection theory, which is a high-level behavioral theory of kin altruism, is a source of endless confusion. For most sociobiologists, kin selection remains, as conceived by Maynard Smith (1964), a social dynamic based on close genealogical association:

> By kin selection I mean the evolution of characteristics which favour the survival of close relatives of the affected individual.

The Wikipedia definition is similar:

> Kin selection is the evolutionary strategy that favours the reproductive success of an organism's relatives, even at a cost to the organism's own survival and reproduction.... Kin selection is an instance of inclusive fitness.

Moreover, while kin selection is a special case of inclusive fitness in the sense that Hamilton's rule applies generally, not just to situations where organisms favor their close genealogical kin, in another sense kin selection is far more general than inclusive fitness. This is because in all but the simplest organisms, kin selection does not describe the behavior at a single locus, but rather an inherently social behavior in which individuals recognize their close relatives through complex phenotypic associations that require higher-level cognitive functioning and synergistic interactions among loci. Indeed, in general these phenotypic associations arise precisely to permit cooperation among close genealogical kin. Moreover, if a mutant gene induces its carriers to interact preferentially with others who have copies of the gene, as recognized by phenotypic characteristics that produced by genes that are linked with the mutant gene, then such a mutant will be evolutionarily successful only if the association is helpful rather than harmful. Then the Inclusive Fitness Harmony Principle implies that this gene will be prosocial. We can thus expect kin selection to arise spontaneously. This of course is observed in many species.

Hamilton's rule and genealogical relatedness

In their critique of inclusive fitness theory, Nowak et al. (2010) remark:

> For many models we find that cooperators are favoured over defectors for weak selection, if a condition holds that is of the form
>
> $$\text{something} > \frac{c}{b}.$$
>
> … an inclusive fitness calculation might derive [this] inequality, but typically that 'something' is not relatedness.

The authors treat this as an empirical observation based on their reading of the literature, but in fact we can show this analytically, closely following Hamilton (1964a).

Consider a social species in which a focal allele imposes a cost on a donor A, but induces A to confer a benefit on a recipient B in the population in some fashion given by the social structure and dynamics of the species. We assume, to simplify the notation, that there is a single recipient, although the argument extends directly to the more general case of several recipients. We also assume that the benefit b and the cost c are the same for all potential recipients, although the argument extends to the case of arbitrary costs and benefits depending on the social identities of donor and recipient.

For each potential recipient B, there is a given probability that B has a copy of the focal allele. Therefore we can partition the population into cells $j = 1, \ldots, k$ such that all individuals in cell j have a copy of the focal allele with probability q_j. Let p_j be the probability that the donor encounters a recipient from cell j, so $\sum_j p_j = 1$. The mean probability that the recipient has a copy of the focal allele is then given by

$$p = \sum_j p_j q_j, \tag{6}$$

We may then use this value of p in Eq. (4).

We can also assume, without loss of generality, that recipients of type j all have genealogical relatedness r_j with the donor. Then any recipient has the focal allele identically by descent with probability r_j and with probability $1 - r_j$ has the focal allele with some probability s_j, so

$$q_j = r_j + (1 - r_j)s_j.$$

If there is no inbreeding, then $s_j = q$ for all j, so $p_j\ (r_j + (1 - r_j)q)$ is the probability that the recipient is of type j and has the focal allele. We then have

$$p = \sum_j p_j\big(r_j + (1 - r_j)q\big) \tag{7}$$

$$= \sum_j p_j r_j (1 - q) + q \tag{8}$$

$$= r + (1 - r)q, \tag{9}$$

as is required by the genealogical interpretation (5). Note that the outbreeding assumption $s_j = q$ is not implied by the assumption of outbreeding for the whole

population, and is likely to fail when genealogical kin have some special social structure, such as being family, clan, or colony members.

It is reasonable to call the array $\{q_j, p_j\}$ the *social structure* of the population with respect to the behavior induced by the focal allele. This array in general is not defined at the level of the focal locus, but at the level of the distribution of the genome in the population, which codes for how individuals seek out specific environments and how, within these environments, they associate assortatively, adopting particular mating patterns, embracing particular rituals and signals, favoring certain patterns of offspring care, and participating in certain forms of social collaboration, in ways giving rise to a particular social structure. Inclusive fitness thus presupposes a general type of social dynamics.

While the simple inequality $br > c$ at first sight appears to connect genealogical relatedness, costs, and benefits at the level of a single locus, in fact a correct derivation of the inequality reveals a complex social structure underlying each of the three terms. This fact does not detract from the importance of Hamilton's rule. Indeed Hamilton's rule must be satisfied by any plausible social structure.

Inclusive fitness without kin selection: an example

This example shows that Hamilton's rule in principle has no necessary relationship with genealogy or kin selection, but rather is an expression of the social structure of the reproductive population. The model is based on Hamilton (1975), which develops a similar model for the same purpose. For related models of positive assortment not based on kin selection see Koella (2000), Nowak (2006b), Pepper (2007), Fletcher and Doebili (2009), and Smaldino et al. (2013).

Consider a population in which groups of size n form in each period. In each group individuals can cooperate by incurring a fitness cost $c > 0$ that bestows a fitness gain b that is shared equally among all group members. Individuals who do not cooperate (defectors) receive the same share of the benefit as cooperators, but do not pay the cost c and do not generate the benefit b. Let p_{cc} be the expected fraction of cooperating neighbors in a group if a focal individual is a cooperator, and let p_{cd} be the expected fraction of cooperating neighbors if the focal individual is a defector. Then the payoff to a cooperator is $\pi_c = bp_{cc} - c$, and the payoff to a defector is $\pi_d = bp_{cd}$.

The condition for the cooperative allele to spread is then $\pi_c - \pi_d = b\,(p_{cc} - p_{cd}) - c > 0$, or

$$b(p_{cc} - p_{cd}) > c. \tag{10}$$

Now p_{cc} is the probability that a cooperator will meet another cooperator in a random interaction in a group, so we can define the relatedness r between individuals, following (5), by

$$p_{cc} = r + (1 - r)q, \tag{11}$$

where q is the mean frequency of cooperation in the population. If we write $p_{dd} = 1 - p_{cd}$ for the probability that a defector meets another defector, then we similarly can write

$$p_{dd} = r + (1 - r)(1 - q), \tag{12}$$

since $1 - q$ is the frequency of defectors in the population. Then we have

$$p_{cc} - p_{cd} = r + (1 - r)q - (1 - (r + (1 - r)(1 - q))) \tag{13}$$

$$= r. \tag{14}$$

Substituting in (10), we recover Hamilton's rule, $br > c$.

Of course, if group formation is random, then $p_{cc} = p_{cd}$ so $r = 0$ and Hamilton's Rule cannot hold. However, to illustrate the importance of social structure, suppose each group is formed by k randomly chosen individuals who then each raises a family of n/k clones of itself. We need not assume parents interact with their offspring, or that siblings interact preferentially with each other. There is no kin selection in the standard sense of Maynard Smith (1964). At maturity, the parents die and the resulting n individuals interact, but do not recognize kin. In this case a cooperator surely has $k - 1$ other cooperators (his sibs) in his group, and the other $n - k$ individuals are cooperators with probability q. Thus

$$p_{cc} = \frac{k-1}{n-1} + \frac{n-k}{n-1}q = q + \frac{(k-1)(q-1)}{n-1}.$$

Similar reasoning, replacing q by $1 - q$ gives

$$p_{dd} = 1 - q + \frac{q(k-1)}{n-1}.$$

Then

$$r = p_{cc} - p_{cd} = p_{cc} - 1 + p_{dd} = \frac{k-1}{n-1},$$

so Hamilton's rule will hold when

$$br = b\left(\frac{k-1}{n-1}\right) > c.$$

Note that the related recipients are all clones of the donor, with relatedness unity, although the r in Hamilton's rule is $(k - 1)/(n - 1)$. The inclusive fitness inequality is accurate here, but kin selection as defined above is inoperative in this model: the altruistic behavior is more likely to spread when the number of families n/k in a group is small.

This model suggests that the interesting question from the point of view of sociobiology is how the genome of the species manages to induce individuals to aggregate in groups of size n and to limit family size to n/k, so that the benefits of cooperation $(b - c)$ can accrue to the population. This is a true miracle of Nature.

What Hamilton's rule does and does not say

Hamilton's inclusive fitness criterion provides an analytically rigorous condition for the evolutionary success of an allele. Because this condition can be satisfied in the

case of genes that sacrifice on behalf of relatives (b, c, $r > 0$). Inclusive information theory contributes to our understanding of social cooperation. However, inclusive fitness theory renders equally plausible biochemical and social behaviors that are comparatively rarely observed. This is because, as is clear from (4) and (5), Hamilton's rule does not place any limitations on the signs of the three variables that occur therein. For instance, assuming random mutations, spiteful behavior, where r, $b < 0$ and $c > 0$, is prima facie equally as likely as r, b, $c > 0$, in the sense that to every pair (r, b) satisfying Hamilton's rule, there is another pair ($-r$, $-b$) that does so as well. Moreover, the condition for the focal allele to enhance the fitness of the population is $n + qn(b - c) > n$, or $b > c$, as we see from (3). If Hamilton's rule is satisfied by (b, r, c) with b, r, $c > 0$, then the rule is equally satisfied by ($-b$, $-r$, c), in which case $b - c < 0$. Inclusive fitness theory therefore makes no prediction concerning the effect of a successful gene on the fitness of its carrier when $c > 0$.

It might be thought that in fact the case $r < 0$ is relatively rarely encountered, and therefore this problem with Hamilton's rule can be safely ignored. The problem, however, is that Hamilton's rule does not explain *why* this case is rarely encountered. Of course, we can simply say that natural selection is unfavorable to such cases, which is true. However, this brings us back to the state of population biology in 1963, before the appearance of Hamilton's rule, when there was no genome-level analysis of social interaction among individuals.

A new set of problems arise if $c < 0$, which is prima facie as plausible as $c > 0$. These problems are suggested by the Inclusive Fitness Disharmony Principle. Indeed, a mutant allele with $c < 0$ has no relatives, so $r = 0$ and Hamilton's rule is necessarily satisfied. Moreover, when the focal allele increases in frequency and $r > 0$, then for $b < 0$ and $|b| > |c|$, the focal allele, which we call *predatory*, is necessarily both harmful ($b < 0$) and antisocial ($b - c < 0$). Evolutionary dynamics of course generally suppress such mutants (Leigh 1977; Frank 2003). The dynamics of free rider suppression, however, are not and cannot be represented in inclusive fitness theory, which does not model the interaction among genetic loci.

In principle there are twelve possibilities for the signs of b, r, and c, ignoring cases where b or c is zero, but four of these cannot satisfy Hamilton's rule. As stressed by Hamilton (1964a), Bourke (2011) and van Veelen (2009), Hamilton's rule allows us to study the other eight categories of social interaction in terms of the concepts *altruistic*, *cooperative*, *predatory*, and *spiteful*. These possibilities are depicted in Table 1. We call the allele *altruistic* if b, $c > 0$, so when Hamilton's rule is satisfied we have $r > 0$ and the focal allele sacrifices an amount of fitness c in order to add an amount b of fitness to the recipient. If $b > 0 > c$, then the allele is *cooperative*, and since $b - c > 0$, the allele contributes unambiguously to the fitness of its carrier. A cooperative allele will always be selected, as in this case Hamilton's rule is always satisfied. If $r < 0$ in this case, the helping allele is conferring benefits preferentially on non-relatives, and so long as these benefits are not too large, i.e., as long as $b|r| < |c|$, the focal allele will increase in frequency. If $c < 0$ and $b < 0$, we call the allele *predatory*. When $r > 0$ in this case, if Hamilton's rule is satisfied provided relatives are not too greatly harmed; i.e., provided $|b|r < |c|$. When $r < 0$ in the selfish case, Hamilton's rule is necessarily satisfied and

Table 1 Variety of behaviors that can satisfy Hamilton's rule

b	c	$r > 0$	$r = 0$	$r < 0$
> 0	> 0	Altruistic	—	
> 0	< 0	Cooperative		
< 0	> 0	—		Spiteful
< 0	< 0	Predatory		

the gene will evolve, although it enhances the fitness of the genome only if the harm b is sufficiently small ($|b| < |c|$). Finally, if $c > 0 > b$, we call the allele *spiteful* because the focal individuals pays a fitness cost to hurt the recipient. According to Hamilton's rule, a spiteful allele will evolve if $r < 0$ and $|b||r| > c$, although it is uniformly destructive to carriers of the genome. (Hamilton 1970, Gardner et al. 2004). For example, a worker in an insect colony who cares for larvae may have a gene that leads it to kill larvae that are not full siblings, or in a diploid species, a germline allele may disable the other allele when the carrier is heterozygous at the focal locus.

A Generalized Hamilton's rule

The focal allele in Hamilton's rule does not confer benefits randomly across the population, but rather targets carriers of the allele (or non-carriers if $r < 0$); i.e., when Hamilton's rule holds, we have $p \neq q$ in (4). But in general the focal allele could be pleiotropic and have more than one targeting pattern. We explore two additional such patterns. First, the focal allele may additionally impose a cost β uniformly on all members of the population, or equivalently, uniformly on all alleles at the focal locus. The case $\beta > 0$ occurs in "tragedy of the commons" cases (Hardin 1968; Wenseleers and Ratnieks 2004), such as when the focal allele bestows benefits preferentially on copies of itself in other carriers through an action that depletes a protein used in chemical processes by somatic cells (Noble 2011). The case $\beta < 0$ occurs in a parasite when the focal allele induces its carriers to suppress an alternative allele at the focal locus that induces carriers to grow so rapidly that it kills its host prematurely (Frank 1996).

We will also consider the case where the focal allele imposes a cost α on all alleles other than the focal allele (Keller and Ross 1998). For example, $\alpha > 0$ can occur if A redirects brooding care from non-relative to relative larvae in an insect colony, and $\alpha < 0$ can occur if the focal allele helps other alleles at the focal locus that benefits carriers by avoiding possibly deleterious homozygosity at the focal locus. We can clearly treat α as cost imposed on all alleles at the focal locus, plus a benefit of equal magnitude enjoyed by carriers of the focal allele. Thus if the population size is n in the current period, population size n' in the next period will include $n + qn(b + \alpha - c)$ individuals because of the behavior induced by the

focal allele, but this will be reduced by $n(\alpha + \beta)q$ due to the effects on non-focal alleles. The number of relatives of the focal allele in the current period is qn, which is increased by the behavior by $qn(pr + \alpha - c)$, and decreased through lower efficiency by $qn(\alpha + \beta)q$. Thus the new population size is given by

$$n' = n(1 - (\alpha + \beta)q) + qn(b + \alpha - c), \tag{15}$$

and (3) becomes

$$\Delta q = \frac{qn(1 - (\alpha + \beta)q) + qn(pb - \alpha - c)}{n(1 - (\alpha + \beta)q) + qnleft(b - \alpha - c)} - q > 0, \tag{16}$$

which simplifies to

$$b(p - q) > (c - \alpha)(1 - q). \tag{17}$$

Substituting $p = r + (1 - r)q$, we get the Generalized Hamilton's rule

$$br > c - \alpha. \tag{18}$$

An alternative derivation of the Generalized Hamilton's rule using the regression approach is presented in Appendix 1.

The effect of an increase in the focal allele on population fitness is the sign of dn'/dq, which is given by

$$\frac{dn'}{dq} = n(b - c - \beta). \tag{19}$$

Note that in the case of Hamilton's rule, which is the above with $\alpha = \beta = 0$, population fitness increases with the frequency of the focal allele in the case of altruism or cooperation, where $b > c$, and decreases in the case of spite ($b - c < 0$). In the case of the Generalized Hamilton's rule, the fitness effect is indeterminate. As we explain below, Hamilton (1964a) included the $\beta \neq 0$ affect in his calculations, but he did not consider the case where the generalized fitness effects are unevenly distributed among the alleles at the focal locus ($\alpha \neq 0$).

It is useful to give descriptive names to the social interactions when α and/or β is nonzero. We may call the case $\alpha > 0$ *theft*, the case $\beta > 0$ *pollution*, and the cases $\beta < 0$ and $\alpha < 0$ as *social generosity* and *non-kin generosity*, respectively. In the socially generous case, we also say that the allele is producing a *public good* (West et al. 2007, p. 57). This follows the common use of the term in economic theory (Olson 1965). Equation (18) shows *the degree of pollution or social generosity has no bearing on whether the allele can evolve*. Moreover, a thieving altruist (b, c, $\alpha > 0$) will evolve, as will a thieving cooperative allele (b,$\alpha > 0 > c$). Finally, the producer of a public good will evolve only if it gains in inclusive fitness from so doing ($br > c$).

The most critical implication of the Generalized Hamilton's rule is that neither social generosity nor pollution has any bearing on whether an allele will evolve, as seen in Eq. (18), despite the fact that a socially generous allele unambiguously enhances the fitness its host, and a polluting allele unambiguously has the opposite effect, as seen in Eq. (19). In addition, a thieving allele does not directly affect the

mean population fitness (see Eq. 19) but it allows the Generalized Hamilton's rule to be satisfied even when $br - c < 0$ (see Eq. 18).

Hamilton's seminal paper (1964a) in fact explicitly includes the pollution and public goods aspect of inclusive fitness, an aspect of his analysis that later writers have ignored. Hamilton called the public good/pollution effect of Generalized Hamilton's rule a *dilution effect* because it affect the rate but not direction of change in the frequency of the focal allele. Hamilton also notes that the dilution effect can lead a successful allele to reduce population fitness. A streamlined presentation of Hamilton's argument, which is quite opaque in the original, is presented in Appendix 3.

The phenotypic gambit

The genome of a multicellular organism includes a myriad of interdependent RNA-producing genes, protein-producing genes and regulatory gene networks. The dynamics of gene interaction are very poorly understood, to the point where it is practically impossible to isolate exactly how a particular gene behaves and interacts with others. Indeed, when we say that a certain allele produces or controls a certain phenotypic trait, what we really mean is that the *absence* of the allele entails the *absence* of the effect. This is, of course, a quite weaker statement, merely asserting that the allele in question contributes in an essential way to the production of the phenotypic effect.

One implication of this state of affairs is that there are few, if any, cases in which a social behavior can be attributed to the choice of an allele at a particular locus of the genome. This fact does not in any way undermine Hamilton's rule, but without additional assumptions, it renders Hamilton's rule inapplicable to analytical models of social behavior without additional assumptions. By far the most plausible and widely used such assumption is the so-called *phenotypic gambit* (Grafen 1984). The phenotypic gambit assumes that a behavior that may be extremely complex at the genetic level can be modeled as though it were the product of the choice of allele at a single locus. In the words of Grafen (1984),

> The phenotypic gambit is to examine the evolutionary basis of a character as if the very simplest genetic system controlled it: as if there were a haploid locus at which each distinct strategy was represented by a distinct allele, as if the payoff rule gave the number of offspring for each allele, and as if enough mutation occurred to allow each strategy the chance to invade.

Quantitative genetics implicitly applies the phenotypic gambit as a basic element in its methodology (Falconer and Mackay 1996). In fact, the haploid assumption is not necessary. There are many examples in the literature of the phenotypic gambit models behavior as controlled by a diploid locus. Moreover, the assumption of a single locus is not necessary, as there is a research tradition in which the production of a phenotypic effect is controlled by two loci, one of which modulates the effects produced at the other locus (Feldman and Liberman 1986). Two-locus models, however, are generally extremely difficult to model and yield few additional insights.

Adopting the phenotypic gambit vastly increases the power of Hamilton's rule, providing a direct analytical relationship between biochemical processes at the level of a single genetic locus and macro-level social behavior. The analytical justification of the phenotypic gambit is, to my knowledge, unknown, but it often produces useful results that are empirically supported. Often, however, it does not. We therefore must treat the phenotypic gambit with a degree circumspection. When a prediction from inclusive fitness theory fails empirically, it is generally due to a faulty phenotypic gambit modeling assumption, not Hamilton's rule itself.

When the interaction of genes is intrinsically frequency dependent, it is sometimes insightful to model two or more behaviors, each of which is assumed to be controlled by a single locus. I given an example of this procedure in Appendix 4, where we model the conflict between queen and worker in controlled the sex ratio of new reproductives.

The myth of inclusive fitness maximization

It asserting that "natural selection leads organisms to become adapted as if to maximize their inclusive fitness," Abbot et al. (2011) doubtless expresses a view with which at least 137 of the world's most prominent population biologists appear to agree. Most of these researchers likely have not personally studied the research on which this assertion is based, but rather submit to the authority of those few who have. The main source cited in support of this statement is a series of papers written by Grafen (1999, 2002, 2006). For instance in a paper devoted to exposing the "misconceptions" of others, West et al. (2011) write:

> Individuals should appear as if they have been designed to maximize their inclusive fitness. Grafen... has formalised this link between the process and purpose of adaptation, by showing the mathematical equivalence between the dynamics of gene frequency change and the purpose represented by an optimisation program which uses an "individual as maximising agent" (IMA) analogy. This emphasises that inclusive fitness is not just an accounting method, it is the component of reproductive success an organism can influence and what organisms should appear to be maximising.

However, Grafen expressedly declares in each of his papers on the subject that additivity across loci, or what is equivalent, frequency independence, is assumed. Others who have carefully studied the conditions under which a population genetics model of gene flow implies fitness maximization at the gene or individual level, including Metz et al. (2008), Gardner and Welsh (2011), and Gardner et al. (2011), require the same assumption. No careful researcher has ever claimed analytical support for the notion that individuals maximize inclusive fitness without making the frequency independence assumption.

If a behavior is adequately captured by the phenotypic gambit applied to a single locus, then frequency independence is trivially satisfied for the model, and the behavior can be modeled as the maximization of inclusive fitness, either of genes or of individuals. But if the phenotypic gambit requires modeling more than one locus,

then the interactions among the modeled loci will generally be frequency dependent. In particular, if the genome's success is based on a pattern of cooperation, promotion, and suppression across loci, which will occur, for instance, if the production of a protein, RNA sequence, or social behavior requires the collaborative activity of many genes (Noble 2011), or if there are frequency dependent social interactions among individuals in a social species (Maynard Smith 1982), then neither genes nor individuals can be characterized as maximizing inclusive fitness.

For a simple example, consider two loci, with alleles a_1,b_1 at the first locus and alleles a_2,b_2 at the second. Suppose b_2 is favored over a_2 by an amount γ_2, but suppose a_1 produces a substance that favors a_2. Suppose a_1 is favored over b_1 by an amount γ_1, but a_2 produces a substance that favors b_1. Let q_1 and q_2 be the frequency of a_1 and a_2 in the population, respectively. Assuming the interactions are linear, we can write

$$\frac{dq_1}{dt} = q_1(\gamma_1 - \beta_1 q_2) \tag{20}$$

$$\frac{dq_2}{dt} = q_2(\beta_2 q_1 - \gamma_2), \tag{21}$$

where γ_1, γ_2, β_1, $\beta_2 > 0$ and γ_1/β_1,$\gamma_2/\beta_2 < 1$. The unique equilibrium of this system is $(q_1^*, q_2^*) = (\gamma_1/\beta_1, \gamma_2/\beta_2)$. This is a neutrally stable focus known as the *Lotka–Voltera* model (Takeuchi 1996). Because paths through time in this model are closed loops, all alleles eventually return to an earlier level of inclusive fitness, which means that the fitness of individual alleles oscillate in time. This precludes fitness maximization (Moran 1964; Akin 1982). Moreover, it is likely that only one of the four combinations (a_i, b_j) for $i,j \in \{1,2\}$ maximizes the fitness of carriers of the genome. Therefore, a mutant regulator gene that suppresses the unfavorable combinations may well arise, in which case it will grow to fixation, as will the favorable alleles at the two loci.

The general point is that if there is a conflict among loci concerning fitness maximization, and if the frequency of alleles at one locus affect the fitness costs and payoffs at other loci, then it is logically impossible that the allele at each locus maximize its inclusive fitness. Rather, the proper setting is *evolutionary game theory*, in which players (i.e., loci) respond strategically to the behavior of other players (i.e., change allele frequency at a locus), and equilibrium, if it is attained, is likely to be a Nash equilibrium in the genome in which each gene plays a best response to the strategy choice of other genes; that is, inclusive fitness cannot be increased at any locus without inducing inclusive fitness-reducing costly counter-moves at other loci (Maynard Smith 1982; Hammerstein and Reichert 1988; Taylor 1989; Weibull 1995; Nowak 2006a; Traulsen and Nowak 2006; Gintis 2009b). In a dynamic context, standard game theory must be supplemented by additional analytical tools that deal with the fact that the rules of the game themselves, as inscribed in the genomes of the players, will evolve according to the dynamics of natural selection (Levin 2009, Akçay and Roughgarden 2011, Akçay and van Cleve 2012). For instance, the rate of recombination in the genome can evolve to counteract collusion among mutant alleles.

Social species are complex adaptive systems

Social species are complex dynamic adaptive systems with several levels of emergent properties (Maynard Smith and Szathmáry 1997; Morowitz 2002). This complexity is due to the nature of the genome, in which there are many interacting particles (loci) with roughly similar structure, but heterogeneous in their details (Holland 1986; Mitchell 2009). Modeling such systems requires sophisticated techniques of which our understanding remains relatively primitive, and modeling promoter and suppressor genes, even in the simplest cases, quickly takes us from population biology to bioengineering and complex social theory.

To say that biological organisms are complex has no explanatory value and does not imply the futility of analytical modeling. Rather, it suggests that understanding multicellular organisms benefits from tools not needed in modeling the behavior of classical dynamical systems. These new tools include laboratory and field experiments, agent-based simulation, evolutionary game theory, and even thick ecological description. In a dynamic context, we must employ analytical tools that recognize the fact that the rules of the game themselves, as inscribed in the genomes of the players, will evolve according to adaptive principles (Gintis 2009a; Levin 2009; Akçay and Roughgarden 2011; Akçay and van Cleve 2012). For instance, the rate of recombination in the genome can evolve to counteract collusion among mutant alleles (Haig and Grafen 1991).

Tension and misunderstanding between the population genetics and systematist communities in sociobiology would dissolve were their respective proponents to recognize the status of social species as complex adaptive systems. Like complex systems in general, sociobiologists must supplement analytical models with high-level synthetic theories of social organization dynamics, as well as biochemical analysis of gene expression, suppression and enhancement.

In their critique of inclusive fitness theory, Nowak et al. (2010) observe that inclusive fitness

> has been the major theoretical attempt to explain the evolution of eusociality. ... We argue that standard natural selection theory in the context of precise models of population structure represents a simpler and superior approach, allows the evaluation of multiple competing hypotheses, and provides an exact framework for interpreting empirical observations.

Their affirmation of the importance of "precise models of population structure" is important and constructive. Precise modeling may even provide a new methodological grounding for sociobiology. Of course, such models must include analytical, agent-based, and high-level synthetic variants, and such models are quite compatible with inclusive fitness theory.

Albert Einstein once was asked why we humans are smart enough to harness atomic energy but not smart enough to contain a potentially catastrophic nuclear arms race. He replied, "It is because politics is more difficult than physics." Metazoan biology is also more difficult than physics. The simplifying assumptions used in physics very often give very close to the right answer, enough so that we can use the laws of physics to create a world of stunning technological complexity. Yet

even the simplest metazoan, with a few hundred genes, is astronomically more complicated than any of our engineering accomplishments.

The core genome as object of selection

A gene generally promotes the fitness of its host only if it is complemented by a well-orchestrated configuration of genes with which it interacts (Sober and Lewontin 1982; Mayr 1997; Hammerstein and Leimar 2006; Noble 2011). This suggests that configurations of genes are the object of selection. Nevertheless, many theorists maintain the contrary position that the gene is the only replicator in a sexually reproducing species (Dawkins 1976). This is not prima facie a very attractive position, but supporters of the gene's eye view maintain that it is forced upon us by the fact that replication in a sexually reproducing organism tears apart and reorganizes higher-order configurations of genes in each generation. Because such configurations do not make copies of themselves, they cannot have fitnesses and cannot be considered either replicators or objects of selection.

This argument was offered by Dawkins (1982), who notes that because in species that reproduce using meiosis and recombination, the genome dies with the body it inhabits. Dawkins concludes that the individual is but a *vehicle* for the transportation of genes across metazoan bodies, writing that a replicator must have a

> low rate of spontaneous, endogenous change, if the selective advantage of its phenotypic effects is to have any significant evolutionary effect.... too long a piece of chromosome will quantitatively disqualify itself as a potential unit of selection, since it will run too high a risk of being split by crossing over in any generation (p. 47).

Perhaps this argument had some cogency in an era before gene frequencing became feasible, but we now know that all but a small fraction of genes in a metazoan species, abstracting from rare mutations, are shared by all individuals in the reproductive population, or fall into a few categories that are preserved across generations. A favorable configuration of genes is thus not torn apart by meiosis and cross-over, but rather evolves to near-fixation in the genome (that is, fixed except for random recurring mutations) or is replaced by an alternative configuration that achieves near-fixation.

Let us define the *core genome* of a sexually reproducing species as a subset of the genome including all loci that have certain key properties ensuring that meiosis and crossover do not interfere with the transmission of fitness-enhancing gene configurations across generations. Included in the core genome are the *fixed loci* and *synonymous loci*. The fixed loci are those in which a single allele is shared by all members of the population, except for rare mutations. The synonymous loci consists of loci in which all alleles, except for low-frequency mutations, produce identical biochemical and phenotypic effects. In addition certain non-synonymous alleles may have fitness neutral, or near-neutral, phenotypic effects (e.g., tail length or eye color). The set of such *fitness neutral gene sets* are stable across generations despite

their somewhat labile internal composition, and are also part of the core genome. For instance, body size may be fitness independent over some range, and many genes interact to produce a phenotypic body size that is generally in the fitness-neutral range. The frequency distribution of these genes in the core genome is determined by natural selection and unchanged by meiosis and crossover.

In addition, if a set of alleles at a particular locus have equal fitness but distinct phenotypic effects, and if this set is preserved across generations, the alleles are likely to be equally fit alternative strategies in a Nash equilibrium among loci, each being a fitness enhancing best response to the probability distribution of the other loci in the genome. We call such alleles *mixed strategy sets*, and we include these in the core genome. For example, a population equilibrium can sustain a positive fraction of altruistic and selfish alleles, or alleles promoting aggressive versus docile behavior, under certain conditions. Similarly, loci that protect carriers against frequency-dependent variations in environmental conditions, including that of bacterial and viral enemies, can be maintained in a polyallelic state as a means of species-level risk reduction. These include the *immune system gene sets* that maintain considerable heterogeneity to deal with a variety of possible infectious agents.

Another example of a mixed strategy gene set is the interaction of suppressor genes and their targets, where the fitness of the suppressor depends on a positive frequency of target genes. Leffler (2013) document such a set stabilized by balancing selection at least since the primate-hominin split. Finally, heterozygote advantage involves a pair of alleles that maintain positive frequency despite the fitness cost to homozygous carriers. We may call this an *overdominance set*. Additional features arise in dealing with sex-linked genes, including maternal-paternal conflict, but these also can be identified as characteristics of the species that are conserved across many generations.

In species that recognize individuals, including many birds and mammals, such recognition is based in part on genes outside the core genome, which are shuffled and redistributed through meiosis and recombination. Species that recognize group differences, such as humans, succeed because of linkage disequilibrium among non-core loci that is maintained by assortative mating.

In sum, the typical phenotypic characteristics of the species, including biochemistry, physiology, and behavioral predispositions, are conserved despite crossover and meiosis. We call the genes that support this phenotypic commonality the species' *core genome* and the loci that are more or less intraspecifically heterogeneous the *variant genome*—see Riley and Lizotte-Waniewski (2009) for an application to bacterial species. The core genome is a replicator and hence subject to the laws of natural selection. Individuals, social groups, as well as their extended phenotypes and constructed niches, are simply vehicles for the expression and evolution of the core genome.

The sociobiological point is that the core genome of social species codes not only for the general phenotypic character of the individual, but also for the species' extended phenotype and predisposes individuals for social behaviors characteristic of the species. The core genome generally remains remarkably stable despite temporal variation in environmental variables, although it is constantly subject to

invasion by successful mutations and reactions to changes in the frequency distribution of predators, prey, parasites, and infectious agents, as well as long-term climate changes.

While the core genome is an object of selection, it is not in any sense a *unit* of selection because it is specified by the frequency distribution of genomes in the population. Moreover, the very notion of units and objects of selection, while perhaps of use for a synthetic understanding of biological evolution, do not appear to play any role in modeling the social structure and dynamics of a reproductive population. However, recognizing the core genome as an object of selection is a useful heuristic in at least two ways. First, while not in any way undermining the insights of the gene's eye view of evolution, it captures the notion that precise combinations of gene interactions are adaptive and hence favored by natural selection. Second, the core genome allows us to conceptualize phenotypic effects that are located not in individuals, but in their social interactions. In other words, the core genome strongly predisposes a social species for certain forms of social behavior, including typical mating patterns, recognized forms of territoriality, and preferred forms of social grouping. The core genome also predisposes organisms to seek out particular natural environments, although there is natural variation in such environments that serve as epigenetic sources of social dynamics and social learning (Galef and Laland 2005; Goodnight et al. 2005; Smaldino et al. 2013).

The core genome is a replicator in the sense of Lewontin (1970). First, mutations in loci of the core genome give rise to *phenotypic heterogeneity*. Second, phenotypic differences can entail *fitness differences among members* of the reproductive population. Finally, such fitness differences are *heritable*. A mutation at a fixed locus, for instance, can lead to increased fitness of carriers of the mutated allele, leading to the increase in frequency of the new allele in the population. The focal locus then drops out of the core genome, but in the long run, with high probability, the mutation will either move to fixation or extinction, restoring the focal locus to the core genome.

Levels of selection

If the species is commonly partitioned into groups, such as insect colonies, schools of fish, herds of elephants, primate communities, and hunter-gatherer groups, then these groupings are the phenotypic expressions of the core genome of the species, as modulated by its characteristic environment, and are subject to selection on that basis.

Another way of expressing this point is to note that sociality in a species is an extension of multicellular cooperation within the genome, and indeed there is a common population model which subsumes both as special and virtually isomorphic cases (Gardner et al. 2007). Because cooperation among loci within the genome is limited to a single instance of a genetic locus, even the simplest of fitness-enhancing processes, such as the sharing of genetic material in microbes and sexual reproduction in metazoans, requires the interaction of two or more copies of key genetic loci, and this can occur only through interactions between carriers. This

interaction of loci is precisely the subject matter of sociobiology, and the processes involved are characteristic of social species. Just as the genome codes for the patterns of interaction among loci in the genome, so it codes for the characteristic patterns of interactions of loci in two or more carriers; i.e., the genome codes for the social structure of the organisms it creates. In particular, social species live in groups precisely when the conditions for such groups are coded in the core genome.

Just as a computer program includes different functional subprograms, some of which work alone and others of which operate with many copies functioning in parallel, with an overarching set of signaling and behavior protocols that facilitate communication and collaboration across subprograms, so the core genome endows its carriers with the protocols for fitness-enhancing interaction among genetic loci located in a plurality of individuals. If the carriers of the core genome form social groups of characteristic size, composition, and behavior, this is because the core genome provides its carriers with instructions as to which environments to seek, and within these environments, how to interact with conspecifics. Social groups are thus genetic rather than epigenetic, although in advanced social species epigenetic cultural forces may be integrally involved in the constitution of social groups. Because social groups are expressions of the core genome, and groups with particular characteristics will evolve only if they enhance the fitness of the reproductive population supported by the core genome, we conclude that if we are to use the term "group selection" at all, it must be understood as section *for the network of genes that underlie the constitution of the social group*.

Conclusion

William Hamilton's inclusive fitness theory is a key analytical tool of modern sociobiology. While inclusive fitness theory justifies selfish gene theory, neither inclusive fitness theory, nor any other plausible theory, supports the notion that genes or individuals in a social species maximize inclusive fitness.

The evolutionary process, from the first RNA molecules to advanced metazoans and complex social species, involves solving the problem of promoting cooperation among selfish genes (Maynard Smith and Szathmáry 1995). That genes generally contribute to the fitness of the individuals in which they reside is the result, not of inclusive fitness maximization, but of a complex evolutionary and intragenomic dynamic involving the suppression of antisocial and promotion of prosocial alleles (Leigh 1971; Buss 1987; Michod 1997; Frank 2003; Noble 2011).

The evolutionary forces that determine the complex interactions among loci in metazoans and among individuals in social species must be studied using, in addition to inclusive fitness theory, the phenotypic gambit (Grafen 1984), evolutionary game theory (Wilson 1977; Taylor 1992, 1996), agent-based modeling (Gintis 2009b), the physiology of suppressor and promoter genes (Leigh 1977; Noble 2011), as well as species-level systematics and ecology.

Acknowledgments I would like to thank Samuel Bowles, David Haig, Steven Frank, Peter Godfrey-Smith, David Queller, Laurent Lehmann, Samir Okasha, Peter Richerson, Joan Roughgarden, Elliot Sober, and Mattijs van Veelen for advice in preparing this paper.

Appendix 1: Regression approach to the generalized Hamilton's rule

Hamilton (1970) developed a more general notion of relatedness based on the Price equation (Price 1970), an approach developed by Queller (1992), Frank (1998), and many others. We can show that this approach can also be used to derive the Generalized Hamilton's rule (suggested by Laurent Lehmann, personal communication).

Suppose in each period each individual is called upon once to help and once potentially to receive help. Only the individual with the helping allele actually helps. Let w_{ij} be the fitness of an individual where $i = 1$ if the individual has the helping allele, $i = 0$ otherwise, and $j = 1$ if the individual is helped, and $j = 0$ if not helped. Then we can write

$$w_{ij} = 1 - ci + bj - (1 - i)\alpha - \beta q, \tag{22}$$

where all parameters are as previously defined. The frequency p_{ij} of type ij is given by

$$\begin{aligned} p_{00} &= (1 - q)(r + (1 - r)(1 - q)) \\ p_{01} &= p_{10} = (1 - q)q(1 - r) \\ p_{11} &= q(r + (1 - r)q). \end{aligned}$$

We can write the regression equations for w_{ij} as

$$\begin{aligned} w_{ij} &= \hat{w}_{ij} + \epsilon_{ij} \\ &= \gamma_0 + \gamma_1 i + \gamma_2 j + \epsilon_{ij}, \end{aligned}$$

where $\hat{w}_{ij}$ is the additive fitness and ϵ_{ij} is an error term. For least squares estimation we minimize

$$\sum_{i=0}^{1}\sum_{j=0}^{1} p_{ij}(w_{ij} - \hat{w}_{ij})^2 = \sum_{i=0}^{1}\sum_{j=0}^{1} p_{ij}\epsilon_{ij}^2$$

with respect to γ_0, γ_1, and γ_2. We find that

$$\begin{aligned} \gamma_0 &= 1 - \alpha - q\beta \\ \gamma_1 &= \alpha - c \\ \gamma_2 &= b. \end{aligned}$$

The additive portion of fitness f_1 for the helping allele and f_0 for the non-helping allele are then given by

$$f_0 = (p_{01}\hat{w}_{01} + p_{00}\hat{w}_{00})/(1-q)$$
$$f_1 = (p_{11}\hat{w}_{11} + p_{10}\hat{w}_{10})/q$$

The condition $f_1 > f_0$ then becomes, after some simplification,

$$br + \alpha > c,$$

which is the correct Generalized Hamilton's rule expression. Note that $br + \alpha - c = \gamma_2 r + \gamma_1$, which means that the reasoning leading to the Generalized Hamilton's rule is the same as for the traditional Hamilton's rule, only the parameters being altered.

Appendix 2: A model of free rider suppression

We can model the suppression of antisocial genes formally in terms of the Generalized Hamilton's rule. Suppose a polluting allele x satisfies (18), so the GHR is simply Hamilton's rule:

$$b_x r_x + \alpha_x > c_x. \tag{23}$$

Suppose further that an allele at locus y imposes a punishment $p_y = -\alpha_y > 0$ on each carrier of the x allele at a cost c_b to itself. The gain to copies of y from suppression of x is

$$b_y = v\alpha_x + q_x\beta_x,$$

where v is the probability that the host of allele y does not contain a copy of the x allele. Allele y also receives the gain b_y so y's net fitness sacrifice is $c_y - b_y$. We assume the cost of punishing p_y is $c_y = \kappa\, p_y$, for some $\kappa > 0$. The punishment p_y is dissipated, so it does not appear as a transfer of fitness to the y allele. Thus the GHR for the y allele becomes

$$(v\alpha_x + q_x\beta_x)(1 + r_y) > \kappa p_y. \tag{24}$$

The effect of suppression on allele x is to increase the cost of x's action from c_x to $c_x + q_y p_y$. Thus the revised GHR for allele x becomes

$$b_x r_x + \alpha_x > c_x + q_y p_y. \tag{25}$$

If there is an interior equilibrium to this two-equation system, it is easy to check that the equilibrium values q_x^* and q_y^*satisfy

$$q_x^* = \frac{p_y q_x \kappa}{(1 + r_y)(v\alpha_x + q_x\beta_x)} \tag{26}$$

$$q_y^* = \frac{b_x r_x - c_x + \alpha_x^*}{p_y}. \tag{27}$$

Clearly if p_y is sufficiently large, the x allele will have a low but non-zero frequency in the population. Thus the GHR shows that an equilibrium level of an antisocial gene can be maintained, and the suppression is more complete the larger

is the punishment p_y, the smaller the cost of punishing $\kappa\ p_y$, the higher the relatedness r_y at the suppressor gene, the stronger the pollution effect β_x, the stronger the thieving effect α_x, and the lower the probability v that the suppressor locus contains a copy of the thieving allele. The GHR also shows that suppression will occur where there are no relatedness effect in the antisocial gene; i.e., when $b_x = 0$ or $r_x = 0$ but $c_x < 0$, so the antisocial gene's gain is purely in personal fitness.

Appendix 3: Hamilton's seminal analysis

The original derivation of the inclusive fitness criterion in Hamilton (1964a) is more general than Hamilton's rule as commonly expressed, and by modifying one of his assumptions, his derivation gives our Generalized Hamilton's rule. Following Hamilton (but with simplified notation), suppose the population is diploid, and the alleles at the focal locus are numbered $k = 1, \ldots, m$. Let R_{ij} be the fitness increment (positive or negative) over baseline fitness unity conferred by an individual of genotype ij at the focal locus on members of the population who carry the i allele, and let T_{ij} be the total fitness increment conferred by an individual with genotype ij on the population. Let the frequency of genotype ij be p_{ij}, and let p_i be the frequency of allele i in the population. With random mating, we have $p_{ij} = p_i p_j$. Hamilton (1964a) assumes this, but this plays no role in the analysis. The total fitness effect of one individual carrying allele i at the focal locus is then given by

$$T_i = \sum_j p_{ij} T_{ij}, \tag{28}$$

and the total fitness effect on allele i due to one carrier of allele i is given by

$$R_i = \sum_j p_{ij} R_{ij}. \tag{29}$$

We then define

$$S_i = T_i - R_i \tag{30}$$

for $i = 1, \ldots, m$, which Hamilton (1964a) calls the *dilution effect* for reasons discussed below. Note that the signs of R_i, S_i, and T_i are indeterminate, but if we assume sufficiently weak selection at this locus, which we do, and if $\overline{T}$ is the total increment in population fitness in one period, then

$$\overline{T} = \sum_i p_i T_i > -1, \tag{31}$$

so population fitness $1 + \overline{T}$ is strictly positive.

Following Hamilton, let us assume that the dilution effect S_i is uniformly distributed over the alleles at the focal locus, and let

$$\overline{S} = \sum_i p_i S_i \tag{32}$$

$$\overline{R} = \sum_i p_i R_i, \tag{33}$$

so $\overline{T} = \overline{R} + \overline{S}$. Then the expression for an increase in the frequency of allele i is given by

$$\Delta p_i = \frac{p_i + p_i R_i + p_i \overline{S}}{1 + \overline{T}} - p_i \tag{34}$$

$$= p_i \frac{1 + R_i + \overline{S}}{1 + \overline{R} + \overline{S}} - p_i \tag{35}$$

$$= p_i \frac{R_i - \overline{R}}{1 + \overline{T}}. \tag{36}$$

From this we get Hamilton's rule, in the form that allele i will increase through natural selection exactly when

$$R_i > \overline{R}. \tag{37}$$

We can thus *define* the inclusive fitness of allele i as R_i. Because the rate growth of allele i is $g_i = \Delta p_i / p_i$, we observe that $R_i > R_j$ exactly when $g_i > g_j$, so alleles at the focal locus are relatively successful in proportion to their inclusive fitness. If the R_i are frequency independent, we can then say that genes "maximize their inclusive fitness."

As Hamilton stresses, the sign and magnitude of $\overline{S}$ do not affect (37), but only the rate at which the frequency of allele i changes in the population. It is for this reason that Hamilton calls $\overline{S}$ a 'dilution' effect.

Hamilton assumes without comment that $R_i > 0$, but this need not be the case. For instance, suppose $m = 2$, $R_{22} = 0$, and $R_{21} = -\alpha < 0$; i.e., there are two alleles, and the heterozygote imposes a fitness loss α on the second allele. Then $R_2 = -\alpha\, p_{12}$, $\overline{R} = -\alpha p_2 p_{12}$, and $R_i - \overline{R} = R_i + \alpha p_2 p_{12}$, which can be positive even if $R_i < 0$. This of course is the case of the antisocial allele.

In general, the assumption that $\overline{S}$ is uniformly distributed among alleles at the focal locus is overly restrictive. In biochemical terms it prevents using inclusive fitness theory to analyze segregation distortion and other allele actions that disfavor other alleles at the focal locus (Ratnieks 1988; Ratnieks and Reeves 1992; Burt and Trivers 2006), or to analyze social helping behaviors (b, r, $c > 0$) that involve imposing costs on non-relatives. So let us assume that allele i at the focal locus receives fraction $\gamma_i \overline{S}$, where $\gamma_i \geq 0$ and $\sum_i \gamma_i = 1$. Then (34) becomes

$$\Delta p_i = \frac{p_i + p_i R_i + p_i \gamma_i \overline{S}}{1 + \overline{T}} - p_i \tag{38}$$

$$= p_i \frac{1 + R_i + \gamma_i \overline{S}}{1 + \overline{R} + \overline{S}} - p_i \tag{39}$$

$$= p_i \frac{R_i - \overline{R} - (1 - \gamma_i)\overline{S}}{1 + \overline{T}} \tag{40}$$

It is clear that the revised condition for allele i to proliferate,

$$R_i - (1 - \gamma_i)\overline{S} > \overline{R}, \tag{41}$$

is the appropriate generalization of our GHR. Note that this inequality can be satisfied even if $R_i < 0$, so the successful allele is uniformly harmful to carriers of the genome.

Note that this analysis does not depend on any particular notion of relatedness. However, Eq. (37) reduces to our expression for Hamilton's rule (1) if we assume there are two alleles at the focal locus one of which is the wild type with zero fitness contribution and the other conferring fitness b on all individuals other than itself, with self-fitness increment $-c$. In this case $R_i = br - c$, $\overline{R} = p_i(br - c)$, so (37) reduces to (1). In addition $T = b - c$ and $S = (1 - r)b$, which are both positive when $b > 0$. This gives the standard contemporary interpretation of Hamilton's rule. In particular, if the gene is indeed a helping gene ($b > 0$), some of the benefits to the recipient will be directed to non-carriers of the allele, so all members of the population gain from the helping behavior. Moreover, we get the GHR expression (18) with $\gamma_i = 0$ and $\overline{S} = -\alpha$.

The dilution effect is important not because it affects the rate of change in the frequency of the focal allele, but because when $\overline{S} < 0$, the success of the focal allele can come at the expense of a lower mean population fitness $1 + \overline{T}$ even when $\overline{R} > 0$. Indeed, the above analysis shows that the conditions for allele success and the conditions for contributing to the success of the reproductive population are distinct.

Appendix 4: Sex allocation and the phenotypic gambit

The following example, based on Charnov (1978), clarifies the meaning of conflict at the gene level in a sociobiological setting, deploying the phenotypic gambit to model both the behavior of the queen and the worker. The model shows how the resolution of this conflict deviates from the model of conflict based on divergence of genetic "interests" and inclusive fitness maximization as the conflict-resolving process. The genetic analysis in this case is sufficiently simple that a complete gene-level treatment is possible with the assumption that either the queen or the workers control the relative allocation of resources devoted to the production of male and female reproductives.

Consider a eusocial haplodiploid species where each colony has one queen, singly mated and mated for life. The workers, all female, raise the colony's brood, which consists of male and female eggs due to queen oviposition, and male eggs due to worker oviposition. Workers cannot produce female zygotes because there are no males in the population except during the mating season.

We assume first that the queen controls the proportion of female and male reproductives, according to a preference that, however complexly regulated in the queen's genome, can be represented by a single locus with alleles a and A, subject to Medelian segregation, the mutant allele A being dominant. We assume an aa queen prefers a female-male sex ratio of r, whereas an Aa or AA queen prefers a sex ratio $\hat{r}$.

We suppose that *aa* and *a* are fixed in the population and investigate the conditions for the *A* allele to invade (Charnov 1978). More precisely, we seek to specify a value of r that cannot be invaded by a distinct sex ratio $\hat{r}$ preferred by a mutant.

We denote the colonies by *xyz*, where $x, y, z \in \{a, A\}$, with *xy* being the queens type and *z* being her mate's type. Because the mutant is rare, we can ignore colonies with more than one mutant type, and we can ignore females of type *AA*. This assumption is purely for ease of exposition, and cannot change the results, provided the mutant is sufficiently rare. The three remaining types are then *aaa*, *Aaa*, and *aaA*. Let n_{xyz} be the number of colonies of type *xyz*.

We assume the colony has one unit of resources to expend on raising reproductives, so an *aa* queen devotes a fraction r resources to gynes (female reproductives) and $1 - r$ to males (all of which are reproductive), while an *Aa* queen devotes $\hat{r}$ and $1 - \hat{r}$ to gynes and males, respectively. We normalize the colony size, which we assume the same for all colonies, to unity. Let s_f and s_m be the expected number of gynes and males, respectively, that survive to reproductive age per unit of resource devoted to their care. We assume all surviving gynes found new colonies, and all surviving males compete equally successfully for mating opportunities. We also assume that in each period a colony has a probability p of persisting until the next period, and workers produce a fraction q of new males.

In the current period, an *aaa* colony produces r gynes of type *aa* and $1 - r$ males of type *a*, so the population produces $n_{aaa} r s_f$ new gynes of type *aa*. An *aaa* colony passes $n_{aaa}(1 - r)s_m$ males of type *a* and $n_{aaa}(p + rs_f)$ gynes of type *aa* to the next generation of reproductives. If ϵ is the fraction of *A* males at mating time, then this gives rise to $n_{aaa}(p + rs_f)(1 - \epsilon)$ colonies of type *aaa* and $n_{aaa} r s_f \epsilon$ colonies of type *aaA*.

The *aaA* colony produces *Aa* gynes, with r resources devoted to gynes, and resources $(1 - r)$ to males, all of which are of type *a*. Thus there are $n_{aaA} r s_f$ new *Aa* gynes and $n_{aaA}(1 - r)s_m$ new males in the population from *aaA* colonies. The queen produces a fraction $(1 - q)$ of the males, all of which are *a*. The workers, which are *Aa*, produce a fraction q of the males, half of which are *a* and half are *A*. Thus *aaA* colonies contribute $n_{aaA}(1 - r)qs_m/2$ males of type *A* and $n_{aaA} r\, s_f$ new *Aa* gynes to the next generation.

The *Aaa* colony produces half *aa* and half *Aa* gynes and a fraction $\hat{r}$ are devoted by the *Aa* queen to new gynes and $(1 - \hat{r})$ to males. The gynes are half *aa* and half *Aa*, while the males are half *a* and half *A*. Thus *Aaa* colonies contribute $n_{Aaa}(p + \hat{r}s_f/2)$ gynes of type *Aa* and $n_{Aaa}(1 - \hat{r})s_m/2$ males of type *A* to the next generation, a fraction q of which are from worker's eggs.

If n' is the next period frequency, we thus have

$$n'_{aaa} = \left(p + rs_f - \epsilon\right) n_{aaa} \tag{42}$$

$$n'_{aaA} = p n_{aaA} + rs_f \epsilon n_{aaa} \tag{43}$$

$$n'_{Aaa} = n_{aaA} r s_f + \left(p + \frac{\hat{r}s_f}{2}\right) n_{Aaa}, \tag{44}$$

where ϵ is the fraction of *aa* new gynes that mate with *A* males.

The number of new males of type a is $n_{aaa}(1-r)s_m$. The new A males consist of those produced in Aaa and aaA colonies. In an Aaa colony, the queen produces a fraction $(1-q)$ of the males, half of which are A, so $n_{Aaa}(1-q)(1-\hat{r})s_m/2$ males are thus produced. The workers produce a fraction q of males, giving

$$n_{Aaa}(2(1-q)+q)(1-\hat{r})s_m/4) = n_{Aaa}(1-\hat{r})(2-q)s_m/4$$

new A males from Aaa colonies. In aaA colonies the queen devotes a fraction r to the production of males, but only the Aa workers, who produce a fraction r of the males, produce A males. This gives $n_{aaA}(1-r)qs_m/2$ type A males.

The ratio of new A males to new a males is then

$$\epsilon = \frac{n_{Aaa}(1-\hat{r})(2-q)s_m/4}{n_{aaa}(1-r)} + \frac{n_{aaA}q(1-r)s_m}{2n_{aaa}(1-r)s_m} \tag{45}$$

$$= \frac{n_{Aaa}(1-\hat{r})(2-q)}{4n_{aaa}(1-r)} + \frac{n_{aaA}q}{2n_{aaa}} \tag{46}$$

We thus have

$$n'_{aaa} = (p + rs_f - \epsilon)n_{aaa} \tag{47}$$

$$n'_{aaA} = \left(p + \frac{rs_f q}{2}\right)n_{aaA} + \frac{rs_f(1-\hat{r})(2-q)}{4(1-r)}n_{Aaa} \tag{48}$$

$$n'_{Aaa} = rs_f n_{aaA} + \left(p + \frac{\hat{r}s_f}{2}\right)n_{Aaa}, \tag{49}$$

Equation (47) shows that when all reproductives are of type aa and a, the population grows at rate $p + rs_f - \epsilon - 1$. The second and third equations are not linked with the first, and show the fate of an invasion of the population by a small number of mutant gynes of type Aa and males of type A. We can depict the dynamics of the invasion by a matrix equation

$$\begin{bmatrix} n'_{aaA} \\ n'_{Aaa} \end{bmatrix} = \begin{pmatrix} p + \frac{rs_f q}{2} & \frac{rs_f(1-\hat{r})(2-q)}{4(1-r)} \\ rs_f & p + \frac{\hat{r}s_f}{2} \end{pmatrix} \begin{bmatrix} n_{aaA} \\ n_{Aaa} \end{bmatrix} \tag{50}$$

The matrix in (50) has positive entries, so it has a maximal real eigenvalue that represents the growth rate of the dynamical system (Elaydi 1999). If the incumbent population is incapable of being invaded, then the value of $\hat{r}$ at which this eigenvalue is maximized must be r.

After some rather tedious calculations, we find that the maximal eigenvalue occurs for $\hat{r} = r$ when r satisfies the equation

$$\frac{(2-q)(1-2r)s_f}{2(3-q)(1-r)} = 0, \tag{51}$$

which implies $r = 1/2$. This prediction of equal investment in males and females, first stated by Fisher (1915) for diploid species, is valid for haplodiploid as well, and does not depend on the fraction q of workers's sons.

The inclusive fitness analysis of this model is much simpler (Trivers and Hare 1976). The queen is related 1/2 to her daughter queens, and each daughter queen

produces one gyne, $(1 - q)$ males, and $q/2$ males *via* her workers, for a total reproductive value of $1 + (1 - q) + q/2 = 1 + (1 + p)/2$. The queen is related $(1 - q) + q/2 = (2 - q)/2$ to a male reproductive, and the male's reproductive value per inseminated queen is the queen herself plus one half the number of males produced by her workers, which is $q/2$. The males reproductive value is thus $(1 + q/2)x$, where x is the ratio of gynes to males in the reproductive population. The population equilibrium occurs when the relatedness times reproductive value for males and gynes are equal:

$$\left(1 - \frac{q}{2}\right)\left(1 + \frac{q}{2}\right)x = \frac{1}{2}\left(2 - \frac{q}{2}\right), \tag{52}$$

giving a sex ratio of

$$x = \frac{4 - q}{4 - q^2}. \tag{53}$$

This differs from the correct ratio of $x = 1$ by at most about 7 %.

Now let us assume that the workers rather than the queen control the allocation of resources to the reproductives. The *aaA* colony produces *Aa* workers, *Aa* gynes, and *a* males. The workers now devote $\hat{r}$ resources to gynes, and $(1 - \hat{r})$ to males. Thus there are $n_{aaA}\hat{r}s_f$ new *Aa* gynes and $n_{aaA}(1 - \hat{r})s_m$ new males in the population from *aaA* colonies. The *aaA* colonies thus contribute $n_{aaA}(1 - \hat{r})qs_m/2$ males of type *A* and $n_{aaA}\hat{r}s_f$ new *Aa* gynes to the next generation. The *Aaa* colony produces half *aa* workers and half *Aa* workers, so the average workers devotes $\bar{r} = (r + \hat{r})/2$ resources to producing gynes and $(1 - \bar{r})$ to producing males. The gynes are half *aa* and half *Aa*, while the males are half *a* and half *A*. Thus *Aaa* colonies contribute $n_{Aaa}(p + \bar{r}s_f/2)$ gynes of type *Aa* and $n_{Aaa}(1 - \bar{r})s_m/2$ males of type *A* to the next generation, a fraction q of which are from worker's eggs.

If n' is the next period frequency, we thus have

$$n'_{aaa} = (p + rs_f)n_{aaa} \tag{54}$$

$$n'_{aaA} = pn_{aaA} + rs_f\epsilon n_{aaa} \tag{55}$$

$$n'_{Aaa} = \left(p + \frac{\bar{r}s_f}{2}\right)n_{Aaa} + n_{aaA}\hat{r}s_f, \tag{56}$$

where ϵ is the fraction of *aa* new gynes that mate with *A* males.

The new *A* males still consist of those produced in *Aaa* and *aaA* colonies. In an *Aaa* colony, the queen produces a fraction $(1 - q)$ of the males, half of which are *A*, so $n_{Aaa}(1 - q)(1 - \bar{r})s_m/2$ are thus produced. The workers produce a fraction q of males, half of the workers are *Aa*, half of their male offspring are *A*, giving

$$n_{Aaa}(2(1 - q) + q)(1 - \bar{r})s_m/4) = (1 - \bar{r})(2 - q)s_m/4$$

new *A* males from *Aaa* colonies. In *aaA* colonies, the workers are *Aa* and half of them produce the fraction q of *A* males, or $n_{aaA}(1 - \hat{r})qs_m/2$ *A* males.

The ratio of new *A* males to new *a* males is then

$$\epsilon = n_{Aaa}\frac{(1-\bar{r})(2-q)/4}{n_{aaa}(1-r)} + \frac{n_{aaA}q}{2n_{aaa}(1-r)} \tag{57}$$

$$= \frac{n_{Aaa}(1-\bar{r})(2-q)}{4n_{aaa}(1-r)} + \frac{n_{aaA}q(1-\hat{r})}{2n_{aaa}(1-r)} \tag{58}$$

We thus have

$$n'_{aaa} = (p + rs_f)n_{aaa} \tag{59}$$

$$n'_{aaA} = \left(p + \frac{rs_f q(1-\hat{r})}{2(1-r)}\right)n_{aaA} + \frac{rs_f(1-\bar{r})(2-q)}{4(1-r)}n_{Aaa} \tag{60}$$

$$n'_{Aaa} = \left(p + \frac{\bar{r}s_f}{2}\right)n_{Aaa} + \hat{r}s_f n_{aaA}, \tag{61}$$

Equation 59 shows that when all reproductives are of type aa and a, the population grows at rate $p + rs_f - 1$. The second and third equations are not linked with the first, and show the fate of an invasion of the population by a small number of mutant gynes of type Aa and males of type A. We can depict the dynamics of the invasion by a matrix equation

$$\begin{bmatrix} n'_{aaA} \\ n'_{Aaa} \end{bmatrix} = \begin{pmatrix} p + \frac{rs_f q(1-\hat{r})}{2(1-r)} & \frac{rs_f(1-\bar{r})(2-q)}{4(1-r)} \\ \hat{r}sf & p + \frac{\bar{r}s_f}{2} \end{pmatrix} \begin{bmatrix} n_{aaA} \\ n_{Aaa} \end{bmatrix} \tag{62}$$

The matrix in (62) has positive entries, so it has a maximal real eigenvalue that represents the growth rate of the dynamical system (Elaydi 1999). If the incumbent population is incapable of being invaded, then the value of $\hat{r}$ at which this eigenvalue is maximized must be r, in which case the eigenvalue must equal, $p + rs_f - 1$, the growth rate of the incumbent population.

After some additional tedious calculations, we find that the maximal eigenvalue occurs for $\hat{r} = r$ when r satisfies the equation

$$\frac{6 - 8r + q(2r-3)}{4(3-q))(1-r))}sf = 0. \tag{63}$$

Solving for the value of r that cannot be invaded, we find

$$r = \frac{6-3q}{8-2q}. \tag{64}$$

This shows that when the queen produces all the males, the worker's desired sex ratio is $r = 3/4$, while if the workers produce the males, the ratio is $r = 1/2$.

The inclusive fitness analysis of this model is similar (Trivers and Hare 1976). The workers are related 3/4 to their sisters, and related $(1 - q)/4 + (3/8)q = (2 + q)/8$ to their brothers. Equation (52) now becomes

$$\left(\frac{2+q}{8}\right)\left(1 + \frac{q}{2}\right)x = \frac{3}{4}\left(2 - \frac{q}{2}\right), \tag{65}$$

giving a sex ratio of

$$r = \frac{1}{1+x} = \frac{3(4-q)}{16+q(1-q)}. \tag{66}$$

This differ from the actual value (64) by at most 2.5 %.

Appendix 5: Hamilton's rule in the diploid case

This section presents a diploid version of the Hamilton's rule. This is often termed the *regression* approach in the literature, but there is in fact no statistical estimation involved in the derivation (Michod and Hamilton 1980). It will be of interest mainly to population biologists.

Consider a reproductive population X with individuals $\{X_i \in X | i = 1, \ldots, n\}$. Suppose the genome has a diploid autosomal locus with two alleles, s (selfish) which leads to a behavior that does not affect the fitness of other individuals, and a (altruistic), which leads its carrier X_i to incur an increased fitness cost c_i over that of the selfish allele, and to bestow fitness benefit b_i distributed over a subset Y_i of recipients. Suppose in addition that the altruistic allele has a social fitness effect β (pollution when $\beta > 0$ or a public good when $\beta < 0$) on both alleles. This cost may be intragenomic, borne by the carrier, or intergenomic, distributed over the population in some arbitrary manner.

Hamilton (1964a) assumes the social fitness effect is distributed uniformly over the genome. This is a significant limitation of his analysis because intragenomically, meiotic drive and other forms of segregation distortion, and socially, altruistic acts that are purchased in part by reducing the fitness of non-relatives, which we may call *thieving effects*, are of extreme importance, although the Inclusive Fitness Harmony Principle suggests that natural selection will limit their observed frequency. We can represent these thieving effects as transfers of fitness α from non-relatives to relatives.

Standard expositions of Hamilton's rule take Y_i to be an individual. This, however, is a restrictive assumption because in many social species individuals interact in groups where it is difficult to apportion the benefit b_i among the various participants. Moreover, as we shall see, Hamilton's rule does not depend on this assumption.

The genotypic value X_g^i of X_i at the focal locus, the frequency of the focal allele at this locus, is 0, 1/2, and 1 for genotypes *ss*, *sa*, and *aa*, respectively. The phenotypic value X_p^i of X_i is 0, h, or 1 according as X_i is *ss* and never confers the benefit, is *sa* and confers the benefit with intensity h, or is *aa* and confers the benefit with intensity one. Here h can have any value, positive or negative, but if the allele effects are additive, then $h = 1/2$. Because there are $2n$ alleles at the focal locus in the population, the frequency of a is $q_a = \sum_i X_g^i/n$. Let Y_g^i be the mean genotype of members of Y_i.

The fitness cost to X_i in the current period is thus $c_i X_p^i$, and the fitness gain to the recipients Y_i is $b_i X_p^i$. The population in the next period is then

$$n(1 - \beta q_a + (b - c)x_p) \tag{67}$$

where $x_p = \sum_i X_p^i/n$ is the mean phenotype of the population, $b = \sum_i b_i X_p^i/x_p$ is the mean benefit, and $c = \sum_i c_i X_p^i/x_p$ is the mean cost. Note that because the thieving effect α is a within-population fitness transfer, it does not appear in (67). The number of donor alleles in the next period is

$$nq_a(1-\beta q_a+\alpha(1-q_a))+\sum_i b_i X_p^i Y_g^i-\sum_i c_i X_p^i X_g^i.$$

The increase in the frequency of the donor allele in the next period, writing the mean genotype of recipients as $q_a^y = \sum_i Y_g^i/n$, is then given by

$$\begin{aligned}
&\frac{nq_a(1-\beta q_a+\alpha(1-q_a))+\sum_i b_i X_p^i Y_g^i-\sum_i c_i X_p^i X_g^i}{n\big(1-\beta q_a+(b-c)x_p\big)}-q_a\\
&=\frac{\Big(\sum_i b_i X_p^i Y_g^i-nbx_p q_a^y\Big)+nq_a\alpha(1-q_a)}{n\big(1-\beta q_a+(b-c)x_p\big)}\\
&-\frac{\Big(\sum_i c_i X_p^i X_g^i-ncx_p q_a\Big)+nbx_p(q_a-q_a^y)}{n\big(1-\beta q_a+(b-c)x_p\big)}\\
&=\frac{\mathrm{cov}(\mathrm{X_p^b},\mathrm{Y_g})-\mathrm{cov}(\mathrm{X_p^c},\mathrm{X_g})+\alpha\mathrm{var}(\mathrm{X_p})-\mathrm{bx_p}(\mathrm{q_a}-\mathrm{q_a^y})}{1-\beta q_a+(b-c)x_p},
\end{aligned}\tag{68}$$

where X_p^b and X_p^c are the variables $b_i X_p^i$ and $c_i X_p^i$, respectively, and X_g is a binomial variable, so $\mathrm{var}(X_p) = nq_a(1-q_a)$. Note that the expression (68) is positive, assuming weak selection, when

$$\frac{\mathrm{cov}(\mathrm{X_p^b},\mathrm{Y_g})+\alpha\mathrm{var}(\mathrm{X_p})-\mathrm{bx_p}(\mathrm{q_a}-\mathrm{q_a^y})}{\mathrm{cov}(X_p^c,X_g)}>1.\tag{69}$$

This inequality is the most general form of Hamilton's rule, including both social fitness and thieving effects. If we assume donors distribute benefits that are, on average, independent from the allelic composition at the focal locus, i.e., $q_a^y = q_a$ then (69) becomes

$$\mathrm{cov}(X_p^b,Y_g)+\alpha\mathrm{var}(X_p)>\mathrm{cov}(X_p^c,X_g).\tag{70}$$

If we further assume that $b_i = b$ and $c_i = c$ for all individuals $i = 1,\ldots,n$, we get the expression:

$$\frac{b\,\mathrm{cov}(\mathrm{X_p},\mathrm{Y_g})+\alpha\mathrm{var}(\mathrm{X_p})}{\mathrm{cov}(X_p,X_g)}>c.\tag{71}$$

Finally, if the effect of the altruistic allele is additive, so $h = 1/2$, then (71) becomes

$$b\frac{\mathrm{cov}(\mathrm{X_p},\mathrm{Y_g})}{\mathrm{var}(X_g)}>c-\alpha.\tag{72}$$

This is a standard expression for Hamilton's rule (Michod and Hamilton 1980), except we have taken into account the thieving effect α (and the pollution/public

good effect β, which does not appear in Hamilton's rule). More generally, for arbitrary h, we have

$$br > cr^p - \alpha, \tag{73}$$

where

$$r = \frac{\mathrm{cov}(\mathrm{X_p}, \mathrm{Y_g})}{\mathrm{var}(X_g)}$$

is the regression coefficient of Y_g on X_p, and r^p is the regression coefficient of X_p on X_g:

$$r^p = \frac{\mathrm{cov}(\mathrm{X_p}, \mathrm{X_g})}{\mathrm{var}(X_g)}.$$

It should be clear that, while we use mathematical terminology from statistical estimation theory, no statistical estimation is in fact involved.

To illustrate the increased generality of the form (70) of Hamilton's rule, suppose the reproductive population is partitioned into social castes $\{Z^j \subset X | j = 1, \ldots, m\}$, where caste j has frequency z_j in the population, and suppose members of the same caste j have the same costs c_j and benefits b_j. Let Y^j be the weighted sum of $\{Y_i | X_i \in Z^j\}$, where each individual is weighted by the number of times the individual appears in the sum. Then we can write (70) as

$$\sum_{j=1}^{m} c_j \Big((b_j \,\mathrm{cov}(Z_p^j, Y_g^j) - c_j \,\mathrm{cov}(Z_p^j, Z_g^j) \Big) + \alpha \,\mathrm{var}(X_p) > 0. \tag{74}$$

Equation (74) shows that in general the social structure of the population allows a caste to be *fundamentally altruistic* in the sense that its net costs of helping exceed the net benefits that the caste contributes to the population. Because the inclusive fitness of caste j is

$$b_j \,\mathrm{cov}(Z_p^j, Y_g^j) - c_j \,\mathrm{cov}(Z_p^j, Z_g^j) < 0 \tag{75}$$

it is then clear that caste j members would maximize their inclusive fitness by simply refusing to contribute to the social process. This shows that in a caste social structure, individuals do not maximize their inclusive fitness. Of course, if castes are genetically determined, then the partition $\{z_j | j = 1, \ldots, m\}$ will be variable across periods and a fundamentally altruistic caste will become extinct in the long run. However, if castes are determined by developmental conditions (e.g., feeding in eusocial insects or socialization in humans), fundamentally altruistic castes can be maintained in the long run.

The sociobiological dynamics of Hamilton's rule

The mapping $X_i \rightarrow Y_i$, which we have taken as given, reflects the *social structure of the reproductive population*. This mapping does not presume any particular set of social relations of kinship, which is why we suggest that *kin selection* is in general an inappropriate description of inclusive fitness dynamics. Note that if the frequency

of the a allele in the population does not affect the fitnesses of alleles at other loci in the genome, then the a allele will move to fixation in the population if Hamilton's rule is satisfied, and will become extinct if the reverse inequality is satisfied. Ultimately, the focal locus will be heterozygous with zero probability.

With frequency dependence, when the focal allele becomes prevalent in the population, if $b - c > 0$, so the allele is beneficial to its carriers, there will be no selection at the level of the genome for genes that suppress the a allele at the focal locus, so the a allele will still move to fixation in the population. When the focal allele is prevalent and $b - c < 0$, there will be natural selection at other loci for genes that either alter the sociobiological mapping $X_i \rightarrow Y_i$ or otherwise suppress the a allele at the focal locus, so that Hamilton's rule no longer holds for the antisocial allele. This is the essence of the Inclusive Fitness Harmony Principle. Of course there may be no likely mutation that suppresses an anti-social a allele, in which case the antisociality reflected in the behavior induced by the a allele will become ubiquitous in the population. natural selection does not guarantee optimality.

This phenomenon also represents a plausible counterexample to Fisher's Fundamental Theorem (Ewens 1969; Price 1972; Frank and Slatkin 1992; Edwards 1994; Frank 1997): as an antisocial allele moves to fixation, the average fitness of population members declines. Some population biologists save Fisher's theorem by calling this a *transmission effect*, and insisting that natural selection always produces fitness-enhancing gene frequency changes (Edwards 1994; Frank 1997; Gardner et al. 2011). This interpretation of natural selection should be avoided because it is arbitrary and difficult to understand for those who are not experts in population biology.

It follows that Hamilton's rule is useful only in charting short-term genetic dynamics. Weak selection and additivity across loci are extremely powerful analytical tools, but in the long run changes in gene frequency at one locus are likely to induce compensatory and synergistic changes at other loci. Indeed, the very mapping $X_i \rightarrow Y_i$ on which Hamilton's rule is based is itself coded in the core genome of the reproductive population, and hence in the long run is modified in the course of evolutionary selection and adaptation.

Altruism among relatives

A relative is a person "allied by blood... a kinsman" (Biology Online). The argument to this point has nothing to do with genealogy, and hence says nothing about altruism among family members. This is an attractive property of our exposition because in a highly social species, individuals interact frequently with non-relatives.

It remains to determine the exact relationship between the sociobiological conception (71) and the genealogical conception of relatedness. We follow Michod and Hamilton (1980), except that we assume the population is outbred at the focal locus. Suppose that each Y_i is an individual recipient, and all recipients have the same genealogical relationship to their donors (e.g., Y_i is a sibling of X_i). Let p_{xyzw} be the joint distribution of genotypes xy for donor and zw for recipient where $x, y, z, w \in \{s, a\}$. Let p^x_{ss}, p^x_{as}, and p^x_{aa} be the marginal distribution of the genotypes

ss, *sa*, and *aa* for the donor (i.e., the fraction of these genotypes in the population), and similarly for p^y_{ss}, p^y_{as}, and p^y_{aa} for the recipient.

We have

$$x_p = hp^x_{as} + p^x_{aa},$$
$$y_p = hp^y_{as} + p^y_{aa},$$

because p^x_{as} is the fraction of *sa* genotypes, their phenotypic value is h, and p_{aa} is the fraction of *aa* genotypes, which have phenotypic value one. Also,

$$p^x_{as} = 2q_n q_a \tag{76}$$

$$p^x_{aa} = q^2_a \tag{77}$$

To derive (76), note that either the paternal allele is *s* with probability $q_n = 1 - q_a$ and the second is *a* with probability q_a, or else the paternal allele is *a* with probability q_a and the second is *s* with probability q_n. The second equation is derived in a similar manner.

We thus have

$$x_p = 2hq_n q_a + q^2_a \tag{78}$$

$$y_p = 2hq_n q_a + q^2_a \tag{79}$$

Note that

$$x_g = 1/2p^x_{as} + p^x_{aa} = q_a$$
$$y_g = 1/2p^y_{as} + p^y_{aa} = q_a.$$

To derive $\mathrm{cov}(X_g, X_p)$, note that

$$\sum_i X^i_p X^i_g/n = hp^x_{as}/2 + p^x_{aa}$$
$$= hq_n q_a + q^2_a$$

Given the values of p^x_{as} and p^x_{aa} from Eqs. (76) and (77), and after algebraic simplification, we find

$$\mathrm{cov}(X_p, X_g) = q_n q_a \alpha/2, \tag{80}$$

where

$$\alpha = 2(h + q_a(1 - 2h). \tag{81}$$

Also,

$$\mathrm{cov}(y_g x_p) = hp_{sasa}/2 + hp_{saaa} + p_{aasa}/2 + p_{aaaa} - y_g x_p.$$

Now let p_{11} be the probability X_i and Y_i share both alleles at the focal locus identically by descent, let p_{10} be the probability the share one allele at the focal locus identically by descent, and let p_{00} be the probability they share neither allele identically by descent. then we have

$$p_{asas} = 2q_n q_a p_{11} + q_n q_a p_{10} + 4q^2_n q^2_a p_{00} \tag{82}$$

$$p_{asaa} = q_a q_n^2 p_{10} + 2q_n q_a^3 p_{00} \tag{83}$$

$$p_{aaas} = q_n q_a^2 p_{10} + 2q_n q_a^3 p_{00} \tag{84}$$

$$p_{aaaa} = q_a^2 p_{11} + q_a^3 p_{10} + q_a^4 p_{00}. \tag{85}$$

If we define f_{XY} as the probability that a random allele in X_i and a random allele in Y_i are identical by descent, then

$$f_{XY} = p_{11}/2 + p_{10}/4. \tag{86}$$

Then a little algebra shows that the r in Hamilton's rule is given by

$$r = \frac{\mathrm{cov}(\mathrm{X_p Y_g})}{\mathrm{cov}(X_p X_g)} = 2f_{XY}. \tag{87}$$

Note that r is then the expected number of copies of the focal allele in the recipient.

Consider, for instance, the case of siblings. The two share the same allele from the father with probability 1/2, and similarly for the mother. therefore $p_{11} = 1/4$, $p_{10} = 1/2$, and $p_{00} = 1/4$. Substituting these values in (82), we get

$$r = \frac{\mathrm{cov}(\mathrm{Y_g}, \mathrm{X_p})}{\mathrm{cov}(X_g, X_p)} = \frac{1}{2}. \tag{88}$$

Thus the sociobiological definition of relatedness and the genealogical definition coincide.

The haploid form of sociobiological relatedness

We now show that the fraction in haploid expression of Hamilton's rule (17) is precisely the sociobiological definition of relatedness. In this case, s is the selfish gene and a is the altruistic gene at the focal locus. The variance of X_g is now

$$\mathrm{var}(X_g) = q_n q_a$$

and

$$\mathrm{cov}(Y_g X_p) = pq_a - q_a^2$$

so

$$\frac{\mathrm{cov}(\mathrm{Y_g X_p})}{\mathrm{var}(X_g)} = \frac{p - q_a}{1 - q_a},$$

which is equivalent to (17).

References

Abbot P et al (2011) Inclusive fitness and eusociality. Nature 471:E1–E4

Akçay E, van Cleve J (2012) Behavioral responses in structured populations pave the way to group optimality. Am Nat 179(2):257–269

Akçay E, Roughgarden J (2011) The evolution of payoff matrices: providing incentives to cooperate. Proc R Soc B 278:2198–2206
Akin E (1982) Cycling in simple genetic systems. J Math Biol 13(3):305–324
Bourke AFG (2011) Principles of social evolution. Oxford University Press, Oxford
Bowles S, Gintis H (2011) A cooperative species: human reciprocity and its evolution. Princeton University Press, Princeton
Brown JL (1974) Alternate routes to sociality in jays—with a theory for the evolution of altruism and communal breeding. Am Zool 14(1):63–80
Burt A, Trivers R (2006) Genes in conflict: the biology of selfish genetic elements. Harvard University Press, Cambridge
Buss LW (1987) The evolution of individuality. Princeton University Press, Princeton
Charnov EL (1978) Sex-ratio selection in eusocial Hymenoptera. Am Nat 112(984):317–326
Crow JF (1954) Breeding structure of populations. II. Effective population number. In: Kempthorne O, Bancroft TA, Gowen JW, Lush JL (eds) Statistics and mathematics in biology, Iowa State University Press, Ames, IA, pp 543–556
Darwin C (1871) The descent of man, and selection in relation to sex. Murray, London
Dawkins R (1976) The selfish gene. Oxford University Press, Oxford
Dawkins R (1982) Replicators and vehicles. In: King's College Sociobiology Group (eds) Current problems in sociobiology. Cambridge University Press, Cambridge, pp 45–64
Dawkins R (1996) The blind watchmaker: why the evidence of evolution reveals a universe without design. W. W. Norton, New York
Dobzhansky T (1953) A review of some fundamental concepts and problems of population genetics. In: Cold Springs Harbor Symposium, pp 1–15
Edwards AWF (1994) The fundamental theorem of natural selection. Biol Rev 69:443–474
Elaydi SN (1999) An introduction to difference equations. Springer, New York
Ewens WJ (1969) A generalized fundamental theorem of natural selection. Genetics 63:531–537
Falconer DS, Mackay TFC (1996) Introduction to quantitative genetics. Benjamin Cummings, New York
Feldman MW, Liberman U (1986) An evolutionary reduction principle for genetic modifiers. Proc Natl Acad Sci 83(13):4824–4827
Fisher RA (1915) The evolution of sexual preference. Eugen Rev 7:184–192
Fletcher JA, Doebili M (1915) A simple and general explanation for the evolution of altruism. Proc R Soc B 276:13–19
Foster KR, Wenseleers T, Ratnieks FLW (2001) Spite: Hamilton's unproven theory. Ann Zool Fennici 38:229–238
Frank SA (1996) Models of parasite virulence. Q Rev Biol 71:37–78
Frank SA (1997) The Price equation, Fisher's fundamental theorem, kin selection, and causal analysis. Evolution 51(6):1712–1729
Frank SA (1998) Foundations of social evolution. Princeton University Press, Princeton
Frank SA (2003) Repression of competition and the evolution of cooperation. Evolution 57:693–705
Frank SA, Slatkin M (1992) Fisher's fundamental theorem of natural selection. Trends Ecol Evol 7(3):92–95
Galef BG, Laland KN (2005) Social learning in animals: empirical studies and theoretical models. Bioscience 55(6):489–499
Gardner A, Welsh JJ (2011) A formal theory of the selfish gene. J Evol Biol 24:1801–1813
Gardner A, West SA, Buckling A (2004) Bacteriocins, spite and virulence. Proc R Soc Lond B 271:1529–1535
Gardner A, West SA, Wild G (2011) The genetical theory of kin selection. J Evol Biol 24:1020–1043
Gardner A, West SA, Barton NH (2007) The relation between multilocus population genetics and social evolution theory. Am Nat 169(2):207–226
Gardner H (2009) The bounds of reason: game theory and the unification of the behavioral sciences. Princeton University Press, Princeton
Gardner H (2009) Game theory evolving. 2nd edn. Princeton University Press, Princeton
Goodnight C, Rauch E, Sayama H, De Aguiar MA, Baranger M, Bar-yam Y (2005) Evolution in spatial predator–prey models and the 'prudent predator': the inadequacy of steady-state organism fitness and the concept of individual and group selection. Complexity 13(5):23–44
Grafen A (1984) Natural selection, kin selection, and group selection. In: Krebs JR, Davies NB (eds) Behavioural ecology: an evolutionary approach, Sinauer, Sunderland, MA

Grafen A (1985) A geometric view of relatedness. In: Dawkins R, Ridley M (eds) Oxford surveys in evolutionary biology, vol 2, Oxford University Press, Oxford, pp 28–89
Grafen A (1999) Formal darwinism, the individual-as-maximizing-agent: analogy, and bet-hedging. Proc R Soc Lond B 266:799–803
Grafen A (2002) A first formal link Between the Price equation and an optimization program. J Theor Biol 217:75–91
Grafen A (2006) Optimization of inclusive fitness. J Theor Biol 238:541–563
Haig D, Grafen A (1991) Genetic scrambling as a defence against meiotic drive. J Theor Biol 153(4):531–558
Hamilton WD (1963) The evolution of altruistic behavior. Am Nat 96:354–356
Hamilton WD (1964a) The genetical evolution of social behavior, I. J Theor Biol 7:1–16
Hamilton WD (1964b) The genetical evolution of social behavior, II. J Theor Biol 7:17–52
Hamilton WD (1970b) Selfish and spiteful behaviour in an evolutionary model. Nature 228:1218–1220
Hamilton WD (1975) Innate social aptitudes of man: an approach from evolutionary genetics. In: Robin F (eds) Biosocial anthropology, Wiley, New York, pp 115–132
Hamilton WD, Peter H, Leimar O (2006) Cooperating for direct fitness benefits. J Evol Biol 19:1400–1402
Hamilton WD, Reichert S (1988) Payoffs and strategies in spider territorial contests: ESS analysis of two ecotypes. Evol Ecol 2:115–138
Hardin G (1968) The tragedy of the commons. Science 162:1243–1248
Holland JH (1986) Escaping brittleness: the possibilities of general purpose learning algorithms applied to parallel Rule-based systems. In: Michalski RS, Carbonell JG, Mitchell TM (eds) Machine learning: an artificial intelligence approach, Morgan Kaufmann, Los Altos, CA
Hölldobler B, Wilson EO (1990) The ants. BelKnap Press, Cambridge
Keller L, Ross KG (1998) Selfish genes: a green beard in the red fire ant. Science 394:573–575
Koella JC (2000) The spatial spread of altruism versus the evolutionary response of egoists. Proc R Soc B 267(1456):1979–1985
Krackauer AH (2005) Kin selection and cooperative courtship in wild turkeys. Nat Biotechnol 434:69–72
Leffler EM (2013) Multiple instances of ancient balancing selection shared between humans and chimpanzees. Science 339:1578–1582
Leigh EG (1971) Adaptation and diversity. Freeman, Cooper, San Francisco, CA
Leigh EG (1977) How does selection reconcile individual advantage with the good of the group? Proc Natl Acad Sci USA 74(10):4542–4546
Levin SA (2009) Games, groups, and the global good. Springer, New York
Lewontin RC (1970) The units of selection. In: Johnston R (eds) Annual review of ecology and systematics, Annual Review Inc., Palo Alto
Malécot G (1948) Les Mathématiques de l'Hérédité. Masson, Paris
Mandeville B (1924) The fable of the bees: private vices, publick benefits. Clarendon, Oxford, 1924[1705]
Mas-Colell A, Whinston MD, Green JR (1995) Microeconomic theory. Oxford University Press, New York
Maynard Smith J (1964) Group selection and kin selection. Nature 201:1145–1147
Maynard Smith J (1982) Evolution and the theory of games. Cambridge University Press, Cambridge
Maynard Smith J (1988) Evolutionary progress and levels of selection. In: Nitecki MH (eds) Evolutionary progress, University of Chicago Press, Chicago, pp 219–230
Maynard Smith J, Szathmáry E (1995) The major evolutionary transitions. Nature 374:227–232
Maynard Smith J, Szathmáry E (1997) The major transitions in evolution. Oxford University Press, Oxford
Maynard Smith J, Ridpath MG (1972) Wife sharing in the Tasmanian native hen, *Tribonyx mortierii*: a case of kin selection? Am Nat 96:447
Mayr E (1997) The objects of selection. Proc Natl Acad Sci 94:2091–2094
Metz JAJ, Mylius SD, Diekmann O (2008) When does evolution optimize. Evol Ecol Res 10:629–654
Michod RE (1997) Cooperation and conflict in the evolution of individuality. 1. The multilevel selection of the organism. Am Nat 149:607–645
Michod RE, Hamilton W (1980) Coefficients of relatedness in sociobiology. Nature 288:694–697
Mitchell M (2009) Complexity: a guided tour. Oxford University Press, Oxford
Moran PAP (1964) On the nonexistence of adaptive topographies. Ann Hum Genet 27:383–393

Morowitz H (2002) The emergence of everything: how the world became complex. Oxford University Press, Oxford
Noble D (2011) Neo-Darwinism, the modern synthesis and selfish genes: are they of use in physiology? J Physiol 589(5):1007–1015
Nowak MA (2006) Evolutionary dynamics: exploring the equations of life. Belknap Press, Cambridge
Nowak MA (2006) Five rules for the evolution of cooperation. Science 314:1560–1563
Nowak MA, Tarnita CE, Wilson EO (2010) The evolution of eusociality. Nature 466(26):1057–1062
Olson M (1965) The logic of collective action: public goods and the theory of groups. Harvard University Press, Cambridge
Pepper JW (2007) Simple models of assortment through environmental feedback. Artif Life 13(1):1–9
Price GR (1970) Selection and covariance. Nature 227:520–521
Price GR (1972) Fisher's 'fundamental theorem' made clear. Ann Hum Genet 36:129–140
Queller DC (1992) A general model for kin selection. Evolution 42(2):376–380
Ratnieks F (1988) Reproductive harmony via mutual policing by workers in eusocial Hymenoptera. Am Nat 132(2):217–236
Ratnieks FLW, Reeves HK (1992) Conflict in single-queen hymenopteran societies: the structure of conflict and processes that reduce conflict in advanced eusocial species. J Theor Biol 158:33–65
Ridley M, Grafen A (1981) Are green beard genes outlaws? Anim Behav 29(3):954–955
Riley MA, Lizotte-Waniewski M (2009) Population genomics and the bacterial species concept. Methods Mol Biol 532:367–377
Rousset F, Billard S (2007) A theoretical basis for measures of kin selection in subdivided populations. Proc Natl Acad Sci 61:2320–2330
Skutch AF (1961) Helpers among birds. Condor 63:198–226
Smaldino PE, Schank JC, McElreath R (2013) Increased costs of cooperation help cooperators in the long run. Am Nat 181(4):451–463
Sober E, Lewontin RC (1982) Artifact, cause, and genic selection. Philos Sci 48:157–180
Takeuchi Y (1996) Global dynamical properties of Lotka–Volterra systems. World Scientific, Singapore
Taylor P (1989) Evolutionary stability in one-parameter models under weak selection. Theor Popul Biol 36(2):125–143
Taylor P (1992) Altruism in viscous populations: an inclusive fitness model. Evol Ecol 6:352–356
Taylor P (1996) Inclusive fitness arguments in genetic models of behavior. J Math Biol 34:654–674
Traulsen A, Nowak MA (2006) Evolution of cooperation by multilevel selection. Proc Natl Acad Sci 103(29):10952–10955
Trivers RL, Hare H (1976) Haplodiploidy and the evolution of social insects. Science 191:249–263
Uyenoyama MK, Feldman MW (1980) Theories of kin and group selection: a population genetics approach. Theor Popul Biol 17:380–414
van Veelen Ma (2009) Group selection, kin selection, altruism, and cooperation: when inclusive fitness is right and when it can be wrong. J Theor Biol 259:589–600
Weibull JW (1995) Evolutionary game theory. MIT Press, Cambridge, MA
Wenseleers T, Ratnieks FLW (2004) Tragedy of the commons in *Melipona bees*. Proc R Soc Lond B 271:S310–S312
West-Eberhard MJ (1975) The evolution of social behavior by kin selection. Q Rev Biol 50:1–33
West SA (2009) Sex allocation. Princeton University Press, Princeton, NJ
West SA, Diggle SP, Buckling A, Gardner A, Griffin AS (2007) The social lives of microbes. Annu Rev Ecol Evol Syst 38:53–77
West S, Mouden CE, Gardner A (2011) Sixteen common misconceptions about the evolution of cooperation in humans. Evol Hum Behav 32(4):231–262
Wheeler WM (1928) The social insects. Harcourt, Brace, New York
Wilson DSl (1977) Structured demes and the evolution of group-advantageous traits. Am Nat 111:157–185
Wilson DS, Pollock GB, Dugatkin LA (1992) Can altruism evolve in purely viscous populations? Evol Ecol 6:331–341