JOURNAL OF Evolutionary Biology

doi: 10.1111/j.1420-9101.2011.02236.x

The genetical theory of kin selection

A. GARDNER*†, S. A. WEST* & G. WILD‡

*Department of Zoology, University of Oxford, Oxford, UK
†Balliol College, University of Oxford, Oxford, UK
‡Department of Applied Mathematics, University of Western Ontario, London, ON, Canada

Keywords:

adaptation;
genetic value;
Hamilton's rule;
inclusive fitness;
least-squares regression;
natural selection;
Price equation;
relatedness;
reproductive value;
social evolution.

Abstract

Natural selection operates both directly, via the impact of a trait upon the individual's own fitness, and indirectly, via the impact of the trait upon the fitness of the individual's genetically related social partners. These effects are often framed in terms of Hamilton's rule, $rb - c > 0$, which provides the central result of social-evolution theory. However, a number of studies have questioned the generality of Hamilton's rule, suggesting that it requires restrictive assumptions. Here, we use Fisher's genetical paradigm to demonstrate the generality of Hamilton's rule and to clarify links between different studies. We show that confusion has arisen owing to researchers misidentifying model parameters with the b and c terms in Hamilton's rule, and misidentifying measures of genotypic similarity or genealogical relationship with the coefficient of genetic relatedness, r. More generally, we emphasize the need to distinguish between general kin-selection theory that forms the foundations of social evolution, and streamlined kin-selection methodology that is used to solve specific problems.

Introduction

Natural selection is the part of evolutionary change in heritable characters that is driven by the differential reproductive success of individual organisms (Darwin, 1859; Fisher, 1930; Price, 1970). It is not the sole factor in evolution, nor is it necessarily the major driver of genetic change. However, it receives much attention because it is the part of evolutionary change that generates biological adaptation, and hence leads organisms to appear designed. Darwin (1859) argued that, as natural selection favours those individuals who achieve the greatest reproductive success, individuals will consequently appear designed to maximize their reproductive success (Darwinian fitness). This link between natural selection (dynamics) and phenotypes optimized for fitness maximization (design) has been formalized by Grafen (2002), for a large class of models that excludes social interaction between relatives.

The only fundamental change to Darwin's theory of adaptation by natural selection has been to take account of social interactions between relatives. Darwin (1859, pp. 236–238) launched this avenue of inquiry by asking how the exquisite adaptations of sterile workers could have evolved in the social insects, given that these individuals enjoy no reproductive success and so cannot be directly acted upon by natural selection. His solution was to suppose that their fertile family members carry heritable tendencies for these traits, in latent form, and may pass these on to future generations. He suggested that natural selection acts upon worker traits indirectly, according to their fitness consequences for reproductive family members. The separation of direct fitness (impact upon personal reproductive success) and indirect fitness (impact upon the reproductive success of genetically similar individuals) was emphasized by Fisher (1930, p. 27), who understood that the latter could drive the evolution of altruistic characters such as aposematic coloration in gregarious insects (Fisher, 1930; pp. 158–159). However, Fisher otherwise neglected such effects in his formalization of Darwinian adaptation (Fisher, 1930, p. 27).

The theory of indirect effects was later fully developed by Hamilton (1963, 1964, 1970), whose analyses suggested that natural selection leads the individual to appear designed to maximize the sum of her direct and

Correspondence: Andy Gardner, Department of Zoology, University of Oxford, South Parks Road, Oxford OX1 3PS, UK.
Tel.: +44 (0) 1865 271 271; Fax: +44 (0) 1865 271 249;
e-mail: andy.gardner@zoo.ox.ac.uk

indirect fitness, or 'inclusive fitness'. The formal link between natural selection and inclusive-fitness maximization has been established by Grafen (2006a), in a large class of models that allows for social interactions between relatives. Hamilton also showed that the principle of inclusive fitness can be encapsulated in a pleasingly simple form, which Charnov (1977) termed 'Hamilton's rule'. This states that altruistic behaviour (Table 1) can be favoured by natural selection if $rb - c > 0$, where c is the fitness cost to the actor, b is the fitness benefit to the recipient, and r is the genetic relatedness of the actor and recipient. Maynard Smith (1964) coined the phrase 'kin selection' to describe the process of natural selection operating through indirect fitness effects.

The theory of kin selection describes how natural selection may be separated into direct and indirect components. Hamilton's rule captures this basic separation, and hence appears to have the same generality and explanatory power as the theory of natural selection itself. Specifically, $-c$ and rb represent, respectively, the direct and indirect fitness consequences of a trait, and apply to any character of interest, not just altruism (i.e. r, b & c could all be positive or negative or zero; Table 1). The explanatory power of this approach has been clearly demonstrated in numerous areas of behavioural and evolutionary research, where it allows clear predictions to be made about social adaptations, which can then be tested empirically (Krebs & Davies, 1993; Bourke & Franks, 1995; Mock & Parker, 1997; Frank, 1998; Haig, 2002; Burt & Trivers, 2006; West, 2009).

However, a number of theoretical studies have questioned the validity and generality of Hamilton's rule. First, it has been suggested that Hamilton's rule requires restrictive assumptions, such as genetic variation segregating at only a single locus, weak selection and absence of gene interactions within and between individuals (Cavalli-Sforza & Feldman, 1978; Charlesworth, 1980; Uyenoyama & Feldman, 1980, 1981; Colwell, 1981; Uyenoyama *et al.*, 1981; Michod, 1982; Queller, 1984, 1985; Avilés, 1993; Bulmer, 1994; Wolf *et al.*, 1998, 1999; Gintis, 2000; Wolf, 2003; Wilson, 2005; Doebeli & Hauert, 2006; Fletcher & Doebeli, 2006, 2009; Fletcher *et al.*, 2006; Fletcher & Zwick, 2006; Killingback *et al.*, 2006; Nowak, 2006; Queller & Strassmann, 2006; Traulsen & Nowak, 2006; Wenseleers, 2006; Wilson & Wilson, 2007; Bijma & Wade, 2008; Goodnight *et al.*, 2008; Traulsen *et al.*, 2008; Wilson, 2008; Nowak *et al.*, 2010a,b; Traulsen, 2010; Table 2). Second, where apparent contradictions of Hamilton's rule have been resolved, practitioners of kin-selection analysis have been accused of redefining the cost, benefit and relatedness terms as a *post hoc* fix (Doebeli & Hauert, 2006; Fletcher *et al.*, 2006; Nowak *et al.*, 2010a,b), and departing from original formulations of kin-selection theory.

Table 1 Classification of social behaviours. After Hamilton (1964, 1970) and West *et al.* (2007a,b).

		Fitness effect for recipient	
		+	−
Fitness effect for actor	+	Mutual benefit	Selfishness
	−	Altruism	Spite

Our aim in this article is to clarify the generality of Hamilton's rule, as a statement of how natural selection drives changes in gene frequencies, and to make explicit the links between different theoretical approaches. In the first section, we employ a genetical approach to natural selection, developed by Fisher and used by Hamilton, to provide a derivation of Hamilton's rule. This shows how selection can be divided into direct and indirect components, and makes explicit exactly what the different terms of Hamilton's rule represent. In subsequent sections, we examine the validity and generality of Hamilton's rule, considering the various complexities that have been suggested to make the rule fail to correctly predict the action of natural selection (summarized in Table 2). For illustrative purposes, we work through two issues in detail: nonadditive fitness effects and altruism among nonkin. Finally, in the Discussion, we show that confusion has arisen owing to a failure to distinguish between the general theory of kin selection that forms the foundations of social evolution vs. the streamlined kin-selection methodologies that are used to solve specific problems.

The genetical theory of kin selection

Natural selection is the part of evolutionary change that results from the differential reproductive success of individual organisms (Darwin, 1859; Fisher, 1930; Price, 1970; Box 1). In particular, it concerns change in the heritable component of the phenotype (g; the individual's 'genetic value' for a trait of interest; Price, 1970; Falconer, 1981; Grafen, 1985a; Box 2), and how this is mediated by differences in individual fitness (v; more generally, 'reproductive value'; Fisher, 1930; Price & Smith, 1972; Charlesworth, 1994; Grafen, 2006b; Box 3). A formal statement of the action of natural selection is provided by Price's (1970) equation. This is a change in the arithmetic average of genetic value between successive generations, and is given by:

$$\Delta_S \mathrm{E}(g) = \mathrm{cov}(v, g), \qquad (1)$$

where v is an individual's relative fitness, g is her genetic value for the character of interest, and $\mathrm{cov}(v,g)$ is the covariance of these two quantities taken over all individuals in the population (derivation provided in Box 1; Price, 1970; see also Robertson, 1966, 1968). This covariance form has been termed the 'secondary theorem of natural selection' (Robertson, 1968). For ease of exposition, we will not consider class structure in this article; the action of natural selection in class-structured models is given a very general treatment by

Table 2 Difficulties that have been suggested to render Hamilton's rule invalid.

#	Suggested difficulty	References	Resolution	References
1	Genetic interactions (dominance, epistasis, synergy)	Cavalli-Sforza & Feldman, 1978; Uyenoyama & Feldman, 1980, 1981; Uyenoyama *et al.*, 1981; Michod, 1982; Bulmer, 1994; Fletcher & Doebeli, 2006; Fletcher *et al.*, 2006; Fletcher & Zwick, 2006; Queller & Strassmann, 2006; Wenseleers, 2006; van Veelen, 2009; Traulsen, 2010; Nowak *et al.*, 2010b	A Hamilton's rule regression analysis of additive effects does not require the absence of nonadditive effects	Queller, 1992a; Frank, 1997a, 1998; Lehmann & Keller, 2006a,b; Gardner *et al.*, 2007
2	Strong selection	Charlesworth, 1980; Bulmer, 1994; Wenseleers, 2006; Traulsen *et al.*, 2008; Traulsen, 2010; Nowak *et al.*, 2010a,b	Weak selection is often assumed for simplicity, but is not required in general	Queller, 1992a; Lehmann & Keller, 2006a,b; Gardner *et al.*, 2007
3	Frequency-dependent selection	Bulmer, 1994; Doebeli & Hauert, 2006; Nowak *et al.*, 2010b; Traulsen, 2010	Costs and benefits may be frequency dependent	Hamilton, 1967; Frank, 1986, 1998; Lehmann & Keller, 2006a,b; Gardner *et al.*, 2007
4	Pure strategies	Wenseleers, 2006	Pure strategies are allowed	Gardner *et al.*, 2007
5	Altruism among nonrelatives	Killingback *et al.*, 2006; Nowak, 2006; Fletcher & Doebeli, 2009; Nowak *et al.*, 2010b	Either the behaviour is not altruistic, or else genetic relatedness has been confused with genealogical relationship/DNA sequence similarity/etc.	Grafen, 2007a,b; Taylor *et al.*, 2007b; Lehmann *et al.*, 2007a,b; this article
6	Conditional strategies	Fletcher & Zwick, 2006; Fletcher *et al.*, 2006	Hamilton's rule readily allows for conditional strategies	Hamilton, 1967; Charnov, 1978; Craig, 1979; Maynard Smith, 1980; Parker, 1989; Crozier, 1992
7	Greenbeards	Queller, 1985	Relatedness is defined with respect to role, not with respect to neighbours	Gardner & West, 2010
8	Evolutionary bifurcations	Fletcher & Doebeli, 2006; Doebeli & Hauert, 2006; Wenseleers, 2006	Hamilton's rule correctly describes the change in average genetic value even in the context of a bifurcation	This article
9	Indirect genetic effects (IGEs)	Wolf *et al.*, 1998, 1999; Wolf, 2003; Bijma & Wade, 2008	IGEs are implicitly allowed for in Hamilton's rule	This article
10	Multi-generation effects	Goodnight *et al.*, 2008	Hamilton's rule captures reproductive value effects (long-term genetic contributions to the future)	Rousset & Ronce, 2004; Alizon & Taylor, 2008; Wild *et al.*, 2009; Lehmann, 2010
11	Group selection	Colwell, 1981; Avilés, 1993; Gintis, 2000; Wilson, 2005; Nowak, 2006; Traulsen & Nowak, 2006; Taylor *et al.*, 2007a,b; Traulsen, 2010; Wilson & Wilson, 2007; Wilson, 2008	The multilevel selection approach is mathematically equivalent to the kin selection approach	Hamilton, 1975; Grafen, 1984, 2006a; Wade, 1985; Frank, 1986, 1995; Queller, 1992b; Gardner *et al.*, 2007; Lehmann *et al.*, 2007a,b; Gardner & Grafen, 2009
12	More than two interacting individuals	Nowak *et al.*, 2010b	Hamilton's rule can be applied across any number of interacting individuals	Hamilton, 1964; West & Gardner, 2010

Grafen (2006b; see also Fisher, 1930; Price & Smith, 1972; Taylor, 1990, 1996; Charlesworth, 1994; Taylor & Frank, 1996; Frank, 1997a, 1998; and Taylor *et al.*, 2007a).

Upon the assumption that heritable variation in the character of interest exists ($\mathrm{var}(g) > 0$), the RHS of equation (1) may be expressed as the product of the additive genetic variance and a multiplying coefficient:

$$\Delta_S \mathrm{E}(g) = \beta_{v,g}\mathrm{var}(g). \quad (2)$$

The coefficient $\beta_{v,g} \equiv \mathrm{cov}(v,g)/\mathrm{var}(g)$ has a useful interpretation: if we were to write down the genetic value (g) and relative fitness (v) of every individual in the population, then the coefficient $\beta_{v,g}$ is equal to the slope of the straight line fitted to these (g,v) data by the method of least squares, i.e.:

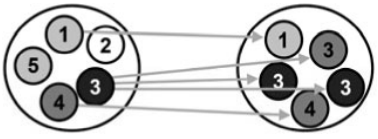
$$\mathrm{E}(v|g) = \mathrm{E}(v) + \beta_{v,g}(g - \mathrm{E}(g)) \quad (3)$$

(see Box 4 for details). The linearity of this least-squares regression is sometimes confused with the assumption of

Box 1 – Selection and transmission

Natural Selection is not Evolution – Fisher (1930, p. vii)

In very general terms, evolutionary change over any specified period of time can be expressed as a sum of selection and transmission components. This is captured by Price's equation, based upon a general mapping between two populations of entities (Price, 1972b, 1995; Frank, 1995). Typically, we suppose that one population is descended from the other, and we denote these 'parent' and 'offspring' populations accordingly. The terms 'parent' and 'offspring' suggest that we are considering evolutionary change over a single generation; however, we are free to consider any two points in time.



To derive Price's equation, we assign every individual in the parent population a unique index $i \in I$, and we assign indices to every individual in the offspring population according to which parent individual they are descended from. When a given individual in the offspring population has more than one ancestor in the parent population (e.g. as in a nonselfing sexual population), each ancestor is awarded their genetic share of the offspring (e.g. mother and father would each be awarded a half share of the offspring, in a diploid population). We denote the 'relative abundance' of the ith parent as q_i, where $\sum_I q_i = 1$. Typically, $q_i = 1/N$, where N is the number of individuals in the parent population, but we are free to assign different individuals different relative abundances, for example according to their relative biomass. Similarly, we denote the relative abundance of the ith parent's descendants in the offspring population as q_i', and this allows us to define the relative fitness of any individual in the parent population as $v_i = q_i'/q_i$ (i.e. the relative growth factor of her immediate lineage). Finally, we can assign each individual in the parent population a value z_i for any character of interest, and we denote the average character value of ith parent's offspring as $z_i' = z_i + \Delta z_i$. The average character value over the parent and offspring populations is $\mathrm{E}(z) = \sum_I q_i z_i$ and $\mathrm{E}(z') = \sum_I q_i' z_i'$, respectively. (Strictly, this is the expectation of the random variable defined as the phenotypic value of an individual drawn at random from the population, with the probability of drawing a particular individual i being equal to her weighting q_i.) Having described the parent and offspring populations, and the mapping between these two, we may write the change in the population average value of the character of interest ($\Delta\mathrm{E}(z) = \mathrm{E}(z') - \mathrm{E}(z)$) as:

$$\Delta\mathrm{E}(z) = \mathrm{cov}(v, z) + \mathrm{E}(v\Delta z), \qquad \text{(B1.1)}$$

where: cov is the covariance and E is the expectation (arithmetic average), each taken over the set of all individuals in the population. The covariance term describes the change that can be ascribed to the statistical association between an individual's character and her relative fitness, and defines selection. (Strictly, this is the covariance of the random variables defined as the phenotypic value and relative fitness of an individual drawn at random from the population, with the probability of drawing a particular individual i being equal to her weighting q_i.) The expectation term describes the change that can be ascribed to character differences between a parent and her offspring, and defines changes associated with transmission.

Natural selection is a particular type of selection that involves genes, the fundamental units of heredity. In Fisher's (1930) scheme, the character of interest is not an individual's phenotype *per se*, but rather her genetic value for any phenotypic character of interest i.e. the heritable portion of the phenotype (Box 2). Moreover, we define change across a single generation. Denoting the genetic value by g_i, then from equation (B1.1) the action of natural selection is given by:

$$\Delta_S\mathrm{E}(g) = \mathrm{cov}(v, g). \qquad \text{(B1.2)}$$

For ease of exposition, we neglect some complicating factors. First, chance differences in individual reproductive success can drive genetic change (random drift). To remove this non-Darwinian evolutionary factor, natural selection is properly defined as an average over future uncertainty (Grafen, 2000). Second, in class-structured populations, one must keep track of differences in offspring class (e.g. male vs. female) as well as offspring number (reviewed by Grafen, 2006b). In such contexts, fitness is properly defined in terms of asymptotic genetic contribution to future generations, i.e. reproductive value (Box 3). In this article, we assume no class structure, for ease of presentation.

linearity in an underlying evolutionary model (see Frank, 1998; Rice, 2004 for more discussion). However, this multiplicative partition of covariance is fully general, and holds irrespective of the degree of nonadditivity of gene-fitness relations. Any impact of dominance, epistasis, etc. upon the immediate action of natural selection is fully accounted for in the least-squares linear regression. Put another way, the goodness of fit of the linear regression model to the population data is irrelevant, because natural selection only 'sees' the additive (i.e. average) effects.

Equation (2) emphasizes that natural selection will occur wherever there is heritable variation ($\mathrm{var}(g) > 0$) in a character that is correlated with fitness ($\beta_{v,g} \neq 0$; Darwin, 1859; Price, 1970). A special case is recovered by taking relative fitness as the character of interest (Fisher, 1941). This gives $\beta_{v,g} = 1$ and hence $\Delta_S \mathrm{E}(g) = \mathrm{var}(g)$ (Frank, 1997a). This is the 'fundamental theorem of natural selection', which states that the change in the average genetic value for relative fitness that can be ascribed to the action of natural selection is equal to the additive genetic variance in relative fitness. As variances are nonnegative, this result captures the directionality of the selective process (Fisher, 1930, 1941; Price, 1972a).

Fisher (1930), following Darwin's (1859) comments on social insects, distinguished two routes for a correlation between an individual's genetic value and her reproductive value. First, an individual's genes (g) can have a direct impact on her fitness, owing to their influence on the phenotype. Second, an individual's genes can have an indirect impact on her fitness, owing to the individual interacting with genetically similar social partners, whose genes (g'; and hence phenotypes) also mediate the focal individual's fitness. We may separate the direct and indirect components of the action of natural selection, using the language of least-squares linear regression (Queller, 1992a; Frank, 1997a, 1998), by imagining that we are fitting the equation:

$$\mathrm{E}(v|g,g') = \mathrm{E}(v) + \beta_{v,g|g'}(g - \mathrm{E}(g)) + \beta_{v,g'|g}(g' - \mathrm{E}(g')) \quad (4)$$

to population data (see Box 4 for details). Substituting this multi-regression model into equation (1) obtains:

$$\begin{aligned} \Delta_S \mathrm{E}(g) &= \beta_{v,g|g'}\mathrm{cov}(g,g) + \beta_{v,g'|g}\mathrm{cov}(g',g) \\ &= \left(\beta_{v,g|g'} + \beta_{v,g'|g} r\right)\mathrm{var}(g), \end{aligned} \quad (5)$$

where $r \equiv \mathrm{cov}(g',g)/\mathrm{cov}(g,g)$ is the genetic relatedness between social partners (Hamilton, 1972; Orlove, 1975; Orlove & Wood, 1978; Michod & Hamilton, 1980; Queller, 1992a; Frank, 1997a, 1998; Box 5). To be clear, the coefficient of genetic relatedness measures the genetic similarity of two individuals, irrespective of why they are similar (e.g. owing to coancestry). For simplicity, we have considered social partners interacting in only one 'role'; a general treatment of potentially multiple roles is provided by Grafen (2006a).

Equation (5) captures the personal (or neighbour-modulated) fitness view of kin selection (Hamilton, 1964; Orlove, 1975; Maynard Smith, 1983; Taylor & Frank, 1996; Frank, 1998). Here, we are concerned with the two routes ($\beta_{v,g|g'}$ vs. $\beta_{v,g'|g}\ r$) by which an individual's genes may impact upon her own personal fitness: through her own expression of these genes, and through her neighbours' expression of copies of the same genes.

Alternatively, we may reassign the indirect fitness effects to the actors responsible for them, giving:

$$\Delta_S \mathrm{E}(g) = \left(\beta_{v,g|g'} + \beta_{v',g|g'} r\right)\mathrm{var}(g). \quad (6)$$

Equation (6) captures the inclusive-fitness view, showing how an individual's genes impact upon her inclusive fitness either by impacting upon her personal fitness ($\beta_{v,g|g'}$; direct fitness effect), or by impacting upon the fitness of individuals to which she is genetically related ($\beta_{v',g|g'}\ r$, where v' is the relative fitness of her social partner; indirect fitness effect) (Hamilton, 1964; Orlove, 1975; Frank, 1998; Rousset, 2004).

The personal-fitness and inclusive-fitness approaches to kin selection are dynamically equivalent: they assign the indirect components of fitness to individuals in different ways, but they predict the same overall response to natural selection when fitness effects are aggregated over the whole population (Hamilton, 1964; Orlove, 1975; Maynard Smith, 1983; Taylor & Frank, 1996; Frank, 1997b, 1998; Rousset, 2004; Grafen, 2006a; Gardner *et al.*, 2007; Taylor *et al.*, 2007a). Hence, irrespective of the approach taken, we may denote the direct effect of an individual's genes on her fitness by $-c = \beta_{v,g|g'}$ and the effect of an individual's genes on her social partner's fitness by $b = \beta_{v,g'|g} = \beta_{v',g|g}$. Consequently, the condition for natural selection to favour an increase in any character of interest ($\Delta_S \mathrm{E}(g) > 0$) is given by Hamilton's rule: $rb - c > 0$.

The assumptions and generality of Hamilton's rule

The previous section has shown that Hamilton's rule of kin selection emerges directly from a general expression of natural selection, simply by partitioning the latter into its direct and indirect components. It would appear to be a very general result. However, within the literature that has grown up around Hamilton's rule, a number of complications have been suggested to limit its generality (summarized in Table 2). We now address each of these in turn.

First, it is often claimed that Hamilton's rule assumes the absence of gene interactions, both within and between individuals (Cavalli-Sforza & Feldman, 1978; Uyenoyama & Feldman, 1980, 1981; Uyenoyama *et al.*, 1981; Michod, 1982; Bulmer, 1994; Fletcher & Doebeli, 2006; Fletcher & Zwick, 2006; Fletcher *et al.*, 2006; Queller & Strassmann, 2006; Wenseleers, 2006; Nowak *et al.*, 2010b; Traulsen, 2010). However, the analysis of the previous section did not require any assumptions to that effect, so this claim is incorrect (see also Frank, 1997a, 1998; Lehmann & Keller, 2006a,b; Gardner *et al.*, 2007). This misunderstanding owes mainly to a confusion of the notion of the additive (average) effect of a gene, on the one hand, with the assumption of no gene

Box 2 – Genetic value

The contribution of imperfectly additive genetic factors divide themselves for statistical purposes into two parts: an additive part which reflects the genetic nature without distortion, and gives rise to the correlations which one obtains; and a residue which acts in much the same way as an arbitrary error introduced into the measurements – Fisher (1918, p. 139)

Natural selection can drive phenotypic change only insofar as organismal characters are heritable. Thus, a proper statement of natural selection concerns changes in the heritable portion of the phenotype. This corresponds to the individual's genetic value, which is the value we predict for her phenotype based upon a least-squares linear regression model that takes as its predictor variables all the genes underlying the character of interest. Formally, the model can be written as:

$$\mathrm{E}(z|x_\gamma \forall \gamma \in \Gamma) = \mathrm{E}(z) + \sum_{\Gamma} \beta_{z,x_\gamma|x_{\gamma'}:\gamma' \neq \gamma,\gamma' \in \Gamma}\left(x_\gamma - \mathrm{E}(x_\gamma)\right), \quad \text{(B2.1)}$$

where: Γ is the set of genic predictors (i.e. gene positions within an individual's genome) impacting upon her phenotype; x_γ is an individual's genic value for the predictor $\gamma \in \Gamma$; $\mathrm{E}(x_\gamma)$ is the population-average genic value for this predictor; z is the individual's phenotypic value; and $\mathrm{E}(z)$ is the population-average phenotypic value. Here, the regression terms β map the set of genic values to phenotypic value, and these can be determined by fitting the model (B2.1) to the set of genic and phenotypic data presented by any population, according to the usual method of least-squares (Box 4). Having fitted the model, an individual's genetic value is simply the phenotypic value that the model predicts for her, based upon her genic data:

$$g = \mathrm{E}(z) + \sum_{\Gamma} \beta_{z,x_\gamma|x_{\gamma'}:\gamma' \neq \gamma,\gamma' \in \Gamma}\left(x_\gamma - \mathrm{E}(x_\gamma)\right). \quad \text{(B2.2)}$$

(Price, 1970; Falconer, 1981; Grafen, 1985a).

interactions, on the other hand (Frank, 1998; pp. 18–19). The additive fitness effect of a gene is an average taken over genotypes, which allows for its dosage in each genotype, using the method of least-squares linear regression (Fisher, 1918). Gene interactions – such as dominance, epistasis and synergy – may mediate the overall additive effect of the gene, and hence impact upon the action of natural selection, but these effects are already taken into account within the regression analysis (they need not be assumed absent). In the context of the linear regression analysis, a nonadditive effect is the discrepancy between the actual and predicted fitness of a genotype and, in any generation, such effects simply do not impact upon the action of natural selection.

Second, it is often claimed that Hamilton's rule breaks down if selection is not vanishingly weak (Charlesworth, 1980; Bulmer, 1994; Wenseleers, 2006; Traulsen *et al.*, 2008; Nowak *et al.*, 2010a,b; Traulsen, 2010). However, the analysis of the previous section allows for selection of arbitrary strength, so again this claim is incorrect (see also Lehmann & Keller, 2006a,b; Gardner *et al.*, 2007). This misconception seems to have arisen for two reasons. In some cases, weak selection can render nonadditive fitness effects negligible, and so (unnecessary) concerns regarding nonadditive effects have sometimes been alleviated by additional assumptions of weak selection. The other reason is that, although relatedness coefficients can be difficult to compute in the context of selection (Bulmer, 1994; but see Ajar, 2003; Gardner *et al.*, 2007; Rousset & Roze, 2007), they are often much easier to compute in the context of neutral populations not undergoing selection. Hence, kin-selection analyses often approximate relatedness coefficients by their neutral counterparts, with the proviso that the approximation is valid only when selection is weak (Taylor, 1996; Taylor & Frank, 1996; Frank, 1998). This simplification is made, just as any other simplifying model assumption, on the basis that it leads to an easier analysis. Furthermore, the assumption of weak selection is often appropriate from an empirical perspective, because many traits are influenced by multiple genes each having only a small effect (Lynch & Walsh, 1998). However, it is incorrect to interpret this simplifying assumption in a specific model as a general failure of Hamilton's rule in the face of strong selection.

Third, it is often claimed that Hamilton's rule breaks down in the context of frequency-dependent selection (Bulmer, 1994; Doebeli & Hauert, 2006; Nowak *et al.*, 2010b; Traulsen, 2010). Moreover, no assumption of frequency independence is built into the analysis of the previous section, so this claim is also incorrect (see also Lehmann & Keller, 2006a,b; Gardner *et al.*, 2007; Ross-Gillespie *et al.*, 2007). A possible reason for this misconception is that Hamilton (1964) framed his principle of inclusive fitness as an extension of Fisher's fundamental theorem of natural selection, which was commonly (though erroneously) assumed to fail in the context of frequency-dependent selection (reviewed by Edwards, 1994). The analysis of the previous section makes clear the generality of the fundamental and secondary theorems of natural selection. Another possible reason for this misconception is the emphasis given by Hamilton (1988) to the frequency independence of kin selection under the assumption of weak selection, which he described as a 'gift from God'. Although frequency independence does allow for an easier kin-selection analysis, frequency-dependent selection does not invalidate Hamilton's rule.

Fourth, Wenseleers (2006) has claimed that Hamilton's rule can only be applied to the evolution of 'mixed' (i.e. probabilistic) strategies, and cannot be applied to the evolution of 'pure' (i.e. hard-wired discrete) strategies. This view appears to arise out of concerns relating to strong selection. High phenotypic variance can potentially translate into high fitness variance between individuals, and this can translate into high fitness variance between alleles (strong selection) if genes are consistently associated with particular phenotypes (pure strategies). By contrast, high fitness variance between individuals may translate into low fitness variance between alleles (weak selection) if genes are not consistently associated with particular phenotypes (mixed strategies). Hence, mixed strategies are a way of imposing weak selection at the gene level even in the context of large fitness differences at the phenotypic level. However, as outlined above, weak selection is not a general requirement for Hamilton's rule to correctly predict the action of natural selection, so this point is moot. Gardner *et al.*, 2007); and this article) provide illustrations of how Hamilton's rule may be applied to models incorporating pure strategies.

Fifth, it is sometimes claimed that altruism can be favoured in the absence of relatedness, implying that Hamilton's rule may fail to correctly predict the action of natural selection (Fletcher & Zwick, 2006; Killingback *et al.*, 2006; Nowak, 2006; Fletcher & Doebeli, 2009). However, in each of the corresponding models, it has subsequently been shown that actors either derive a direct benefit from their behaviour (i.e. the behaviour is not really altruistic), or that they are in fact related to the recipients of their behaviour (relatedness is not really absent) (Lehmann & Keller, 2006a,b; Grafen, 2007a,b; Taylor *et al.*, 2007b; West *et al.*, 2007a,b, 2011; this article). Considering the latter case, relatedness is missed for two broad reasons. One reason is that authors sometimes assume that the lack of an explicit relatedness parameter, or lack of kin discrimination, means that relatedness is not important (e.g. Killingback *et al.*, 2006; Nowak, 2006; Nowak *et al.*, 2010b). In such cases, more formal analyses have demonstrated that altruism is driven by positive relatedness, often because limited dispersal leads to individuals interacting with their close relatives (Grafen, 1984, 2007a,b; Lehmann & Keller, 2006a,b; Lehmann *et al.*, 2007a,b; Taylor *et al.*, 2007b). Another reason that genetic relatedness can be missed is when it is assumed that it requires DNA sequence similarity (e.g. Fletcher & Doebeli, 2009). However, as discussed above, genetic relatedness measures the statistical association between social partners' heritable tendencies for a trait of interest, which does not require that they share the same DNA sequence. The related point that genetical relatedness is distinct from genealogical relationship was first made by Hamilton (1964, 1970, 1975), and has since been analysed in a number of 'greenbeard' models (Dawkins, 1976, 1982; Ridley & Grafen, 1981; Lehmann & Keller, 2006a,b; Gardner & West, 2010; West & Gardner, 2010).

Sixth, Fletcher and colleagues (Fletcher & Zwick, 2006; Fletcher *et al.*, 2006) have suggested that Hamilton's rule cannot capture conditional strategies. However, there is a long history of Hamilton's rule being successfully applied to facultative behaviours (Hamilton, 1967; Charnov, 1978; Craig, 1979; Maynard Smith, 1980; Parker, 1989; Crozier, 1992), so it is difficult to see how this erroneous view has arisen. Moreover, this error is readily demonstrated by the field of sex allocation, which provides a huge theoretical and empirical literature illustrating how kin-selection theory can be used to explain and predict conditional behaviour (Charnov, 1982; Hardy, 2002; West, 2009).

Seventh, Queller (1984) suggested that Hamilton's rule fails to correctly predict the action of natural selection when greenbeard genes underlie social behaviour. These are genes that identify copies of themselves carried in other individuals, and cause their bearer to behave in ways that preferentially benefit carriers of the gene (Hamilton, 1964, 1971, 1975; Dawkins, 1976; Gardner & West, 2010; McGlothlin *et al.*, 2010). Greenbeards can be favoured even in fully mixed infinite populations, whereby neighbouring individuals are not related ($r = 0$). Hence, Queller (1984) suggested that there are no regular indirect effects ($rb = 0$), and added 'correction' terms into Hamilton's rule to capture the greenbeard effect. However, relatedness to neighbours is not the appropriate coefficient of relatedness to use in this context. Because greenbeard individuals can identify which of their neighbours carry the gene, and which do not, then formally we should consider that there are two types of neighbour – carriers and noncarriers – and we should calculate relatedness coefficients appropriate for individuals in each role, separately. As there is a positive relatedness between the greenbeard actor and his bearded neighbours ($r > 0$), and as the actor provides a fitness benefit for these bearded neighbours ($b > 0$), there is an indirect fitness benefit ($rb > 0$), and hence no correction term is needed in order for Hamilton's rule to capture the greenbeard effect (Gardner & West, 2010).

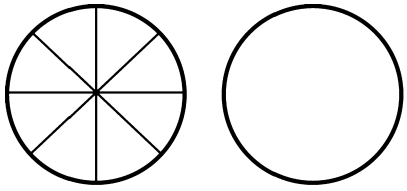
Eighth, there is a concern that Hamilton's rule fails to predict evolutionary dynamics in the context of evolutionary bifurcations, where disruptive selection drives a bimodal distribution of phenotypes (Doebeli & Hauert, 2006; Fletcher & Doebeli, 2006; Wenseleers, 2006). For example, Doebeli & Hauert (2006) have considered that, at a bifurcation point with respect to level of cooperation, there may be no directional selection and mutants exhibiting either a higher or lower level of cooperation may simultaneously invade from rarity. They point out that, in such circumstances, Hamilton's rule is either satisfied (cooperation favoured) or not satisfied (cooperation not favoured), and they suggest that this is inconsistent with both more-cooperative and less-cooperative mutants being able to invade. However,

Box 3 – Reproductive value

To what extent will persons of this age, on average, contribute to the ancestry of future generations? The question is of some interest, as the direct action of Natural Selection must be proportional to this contribution – Fisher (1930, p. 27)

The reproductive value of an individual expresses, in relative terms, the number of copies of that individual's genes expected to be found in the population in the very distant future (Taylor, 1996). Hence, it corresponds to the individual's expected asymptotic share of the future gene pool (Fisher, 1930; Price & Smith, 1972; Grafen, 2006b). Genes carried by individuals with higher reproductive value are expected to be more prevalent in the future population than those carried by individuals with lower reproductive value. When all individuals are equivalent (no class structure), then an individual's (expected) reproductive value is equivalent to her relative fitness, i.e. number of offspring expressed relative to the average. When individuals are not equivalent (class structure) then it is not possible to simply count offspring, as offspring may vary in their quality (i.e. class). Here, reproductive value provides an exchange rate that translates all offspring into a universal currency.

As an example of no class structure, consider a population made up of N asexual individuals. If this population is closed to migration, then the ancestors of its future inhabitants must be among the group of N individuals currently found there. In other words, 100% of the genetic material comprising the future gene pool (i.e. 100% of the 'future success') must have originated from a gene carried by at least one of the population's N present-day residents. Determining the reproductive value of the N present-day residents, then, amounts to determining how to distribute the 'future success' among them. In the absence of selection, the reproductive value of each individual is 1: every individual has average fitness. Alternatively, if one individual monopolizes the reproduction, then its reproductive value is N and the reproductive value of all other individuals is zero. In general, the long-term dynamics of a population determine the reproductive value of each individual.



No selection: 100 % of future success is shared equally among N present-day residents.

Selection: 100 % of future success is awarded to one individual if it monopolizes reproduction.

In a class-structured population, individuals are grouped according to particular qualities they share. For example, a class-structured population could consist of a class that contains only males and another class that contains only females. A class-structured population might alternatively consist of a continuum of classes that corresponds to a continuous range of ages. Because of the intrinsic differences among classes, genetic change in one class may not have the same evolutionary consequences as the same amount of genetic change in another class. To understand how class-specific genetic change reflects evolutionary success in the long term, these class-specific changes must be considered alongside the total reproductive value of the individuals in that class (i.e. the total share of the future gene pool given to that class). In particular, multiplying the total reproductive value of a class by a class-specific gene frequency change, converts the latter into a change in the total share of the future gene pool awarded to the gene in question (Taylor, 1989; Rousset, 2004). Summing all such changes in future shares yields an overall measure of the change in the long-term, evolutionary success of the gene.

Hamilton's rule, as a statement of the action of natural selection, concerns the average genetic change of a population, not the invasion success of rare mutants *per se*, so there is no contradiction here.

Ninth, it has been suggested that Hamilton's rule neglects indirect genetic effects (IGEs), which occur when genes carried by one individual have phenotypic consequences for other individuals (Wolf *et al.*, 1998; Wolf, 2003). Thus, analysis of IGEs has been suggested to generalize beyond a Hamilton's-rule analysis (Wolf *et al.*, 1998; Bijma & Wade, 2008). However, the analysis of the previous section provides an exact and general derivation of Hamilton's rule, based upon the relationship between an individual's genes and her relative fitness, with all phenotypic effects remaining implicit. Hence, the terms of Hamilton's rule already include the impact of IGEs, and no additional terms are needed to account for these. Note that we are not claiming that IGEs are unimportant, just that their existence does not render Hamilton's rule invalid. IGEs can be interesting in their own right, and it will often be helpful to make them explicit within theoretical models, and use these as a means to make broader connections between the theories of kin selection and quantitative genetics, or to make specific predictions for certain situations (Cheverud, 1984; Wolf *et al.*, 1998, 1999; Wolf, 2003; Bijma & Wade, 2008; McGlothlin *et al.*, 2010).

Tenth, it has been suggested that social effects that accrue over multiple generations after the original behaviour was enacted are not capable of being captured

by Hamilton's rule (Goodnight *et al.*, 2008). For the sake of clarity, the analysis of the previous section focuses on a single generation of genetic change and no class structure (Box 3). However, kin-selection analysis readily applies to such problems. One can either consider the impact upon reproductive success of recipients in future generations (Lehmann, 2010) or, equivalently, the changing reproductive values (long-term genetic legacy) of their ancestors in the present generation (Rousset & Ronce, 2004; Alizon & Taylor, 2008; Wild *et al.*, 2009).

Eleventh, a recurring claim is that Hamilton's rule fails to capture the effects of group selection. Such claims usually manifest as suggestions that kin selection cannot account for various phenomena arising in the context of multilevel selection (Colwell, 1981; Avilés, 1993; Gintis, 2000; Wilson, 2005; Nowak, 2006; Traulsen & Nowak, 2006; Taylor *et al.*, 2007a,b; Traulsen, 2010; Wilson & Wilson, 2007; Wilson, 2008; reviewed by West *et al.*, 2008, 2011). Some authors have also attempted to correct Hamilton's rule by adding an additional 'group selection' term (Wilson, 2005; Bijma & Wade, 2008). However, it has long been understood that the kin selection and multilevel (group) selection approaches to social evolution are mathematically equivalent, and merely represent different partitions of the same evolutionary process (i.e. natural selection; Hamilton, 1975; Grafen, 1984, 2006a; Wade, 1985; Frank, 1986, 1995; Queller, 1992b; Rousset, 2004; Gardner *et al.*, 2007; Lehmann *et al.*, 2007b; Gardner & Grafen, 2009). No model of multilevel selection has ever delivered a (correct) prediction that could not be reformulated in terms of kin selection – despite repeated claims to the contrary.

Twelfth, it has been claimed that Hamilton's rule requires pairwise interactions (Nowak *et al.*, 2010b). However, pairwise interactions have not been assumed since the earliest developments of kin-selection theory: Hamilton (1964) made clear that inclusive fitness is defined by summing effects over potentially multiple recipients, and the rule is that natural selection will favour a trait when the total inclusive-fitness effect is positive. For example, if an altruistic act incurs a fitness cost c for the actor and provides a benefit b_1 to one relative and a benefit b_2 to another relative, then the condition for natural selection to favour an increase in altruism is $r_1 b_1 + r_2 b_2 - c > 0$, where r_1 and r_2 are the relatedness of the actor to the former and latter relatives, respectively. For this reason, the condition for altruism to be favoured need not be reducible to the form $b/c > 1/r$, as there may be multiple benefit terms and multiple relatedness terms involved (e.g. see the supplementary online material of West & Gardner, 2010). More generally, kin-selection models of sex allocation and dispersal evolution have long considered fitness consequences for potentially large numbers of recipients (Frank, 1998).

In the next two sections of this article, we consider several of these issues in more detail, by focusing upon two recently proposed models of social evolution in which Hamilton's rule has been argued to fail to correctly predict the action of natural selection. First, we consider games of synergistic cooperation discussed by van Veelen (2009). Our genetical analysis of these models demonstrates that Hamilton's rule correctly predicts the action of natural selection, despite perceived complications of nonadditive fitness effects, strong selection, frequency dependence and pure strategies. We reveal that suggestions to the contrary owe to incorrect computation of cost and benefit terms in Hamilton's rule. Second, we consider Fletcher & Doebeli's (2009) model of altruism among nonkin. Our genetical analysis of this model demonstrates that Hamilton's rule again correctly predicts the action of natural selection, and emphasizes that there may be genetic relatedness even in the absence of genealogical relationship or DNA-sequence similarity. In particular, we reveal that Fletcher & Doebeli's (2009) suggestion to the contrary owes to their incorrect computation of the relatedness term of Hamilton's rule.

Hamilton's rule with synergy

In the previous sections, we provided a derivation of Hamilton's rule, and discussed the extent to which this relies upon restrictive assumptions. In this section, we give a detailed analysis of the issue of nonadditive gene interaction, as the idea that this invalidates Hamilton's rule is an entrenched misconception that continues to manifest in the literature. As a specific example, we consider van Veelen's (2009) claim that Hamilton's rule fails to correctly predict the action of natural selection in the context of three-player evolutionary games with synergistic interaction between the genes of social partners. Here we reanalyse van Veelen's (2009) model, recovering his results, and we show that these are not really in conflict with Hamilton's rule. For completeness, we first analyse the case of two-player games, considering a model developed by Queller (1985).

Two-player games

Queller (1985) suggested that synergistic interaction between social partners could not be captured within the canonical Hamilton's rule, and suggested that an additional correction factor is needed to correctly predict genetic change. The following model is based upon his analysis of a two-player game with synergy. On a historical note, although Queller (1992a) later emphasized the generality of a genetical Hamilton's rule – i.e. our equations (5) and (6) – others have continued to argue that Queller's (1985) result is more general than Hamilton's rule and/or that Hamilton's rule breaks down with synergistic interactions (e.g. Fletcher & Zwick, 2006; van Veelen, 2009).

We assume an infinite population of haploid individuals engaged in two-player games. A single locus controls

Box 4 – Least-squares regression

Nature discovered least squares long before Gauss – Crow (2008)

Suppose we have a set of empirical observations relating a predictor variable x to a target variable y, summarized in a data set (x_i,y_i). We may assume a functional correspondence between x and y, denoted by $y = f(x, (\alpha_j))$, where (α_j) denotes a set of parameters. How are we to choose parameter values for the functional form, in a way that is informed by our empirical observations? In other words, how should we fit the functional relation to our data?

If there is only a single data point (x_1,y_1), then a natural approach is to determine the set of parameter values that minimize the quantity $\epsilon_1 = y_1 - f(x_1, (\alpha_j))$, i.e. the discrepancy (distance) between the predicted and observed values of y. Assuming differentiability, this set $(\alpha_j)^*$ of parameter values must satisfy:

$$\left.\frac{\partial \epsilon_1}{\partial \alpha_k}\right|_{(\alpha_j)=(\alpha_j)*} = 0, \tag{B4.1}$$

and:

$$\left.\frac{\partial^2 \epsilon_1}{\partial \alpha_k^2}\right|_{(\alpha_j)=(\alpha_j)*} > 0, \tag{B4.2}$$

for all k.

If there are several data points, then the parameterization that minimizes the discrepancy for the first data point may not coincide with the parameterizations that minimize the discrepancies for the other data points. What aggregate of the discrepancies for the whole set of n data should we choose as our minimand?

Plotting the discrepancy relating to each data point as a separate, orthogonal axis in n-dimensional space, we can express the overall discrepancy for a given parameterization (α_j) as a single point in that space. A natural metric for the aggregate discrepancy is the Euclidean distance of this point from the origin (a distance of zero representing a perfect fit). From Pythagoras' theorem, this Euclidean distance is given by:

$$\tilde{\epsilon} = \sqrt{\sum_i \epsilon_i^2}. \tag{B4.3}$$

Minimizing the aggregate discrepancy $\tilde{\epsilon}$ is equivalent to minimizing the sum of the squared discrepancies in each dimension $S = \sum_i \epsilon_i^2$. This leads to the following conditions for the best (least-squares) fit parameterization $(\alpha_j)^*$:

$$\left.\frac{\partial S}{\partial \alpha_k}\right|_{(\alpha_j)=(\alpha_j)*} = 0, \tag{B4.4}$$

and:

$$\left.\frac{\partial^2 S}{\partial \alpha_k^2}\right|_{(\alpha_j)=(\alpha_j)*} > 0. \tag{B4.5}$$

The invention of this least-squares method of fitting is credited to Carl Friedrich Gauss. Sheynin (1993) provides an overview of the historical development of this method.

The simplest functional form of any explanatory value is the equation of the straight line: $\mathrm{E}(y|x) = m\,x + c$. This has two parameters: an intercept c and a slope m. For every data point, the discrepancy is of the form $\epsilon_i = y_i - \mathrm{E}(y_i|x_i) = y_i - m\ x_i - c$, and hence the sum of the squared discrepancies is:

$$S = \sum_i (y_i - mx_i - c)^2. \tag{B4.6}$$

The intercept parameter that minimizes the sum of the squares is found by differentiating the RHS of (B4.6) by c, setting the result equal to zero, and solving for c. This obtains:

$$c = \mathrm{E}(y) - m\mathrm{E}(x). \tag{B4.7}$$

Hence, the least-squares linear regression satisfies $\mathrm{E}(y|x) = mx - (\mathrm{E}(y) - m\mathrm{E}(x))$ or, equivalently, $\mathrm{E}(y|x) - \mathrm{E}(y) = m(x - \mathrm{E}(x))$. Note that the best-fit equation passes through the centre of the data: $\mathrm{E}(y|x) = \mathrm{E}(y)$ when $x = \mathrm{E}(x)$.

With this in mind, the sum of squares is given by:

$$S = \sum_i (y_i - \mathrm{E}(y) - m(x_i - \mathrm{E}(x)))^2. \tag{B4.8}$$

The slope parameter that minimizes the sum of squares is found by differentiating the RHS of (B4.8) by m, setting the result equal to zero, and solving for m. This obtains:

$$m = \frac{\mathrm{cov}(x,y)}{\mathrm{var}(x)}. \tag{B4.9}$$

Hence, the ratio of the covariance of two random variables x and y and the variance in x is equal to the slope of the least-squares linear regression line fitted through these (x,y) data. In the main text, we have denoted this ratio $\beta_{y,x}$, and here we have shown the relation to least-squares linear regression.

The least-squares linear regression approach is readily extended to multiple predictors (x_j). Here, the simplest functional form is $\mathrm{E}(y|(x_j)) = \sum_j m_{ij}x_{ij} + c$, and the sum of squares is given by:

$$S = \sum_i \left(y_i - \sum_j m_j x_{ij} - c \right)^2. \tag{B4.10}$$

Following the same procedure as above, we can differentiate the RHS of (B4.10) by c, set the result equal to zero and solve for c to obtain the best-fit intercept parameter:

Box 4 (continued)

$$c = \mathrm{E}(y) - \sum_j m_j \mathrm{E}(x_j). \qquad \text{(B4.11)}$$

Hence, the least-squares regression line satisfies $\mathrm{E}(y|\{x_j\}) = \mathrm{E}(y) + \sum_j m_j(x_j - \mathrm{E}(x_j))$. Note that, as above, this line passes through the centre of the data: $\mathrm{E}(y|\{x_j\}) = \mathrm{E}(y)$ when $x_j = \mathrm{E}(x_j)$ for every j.

With this in mind, the sum of squares is given by:

$$S = \sum\nolimits_i \left(y_i - \mathrm{E}(y) - \sum_j m_j (x_{ij} - \mathrm{E}(x_j)) \right)^2. \qquad \text{(B4.12)}$$

The best-fit value for each slope parameter can be obtained in the usual way. Differentiating the RHS of (B4.12) by m_k, setting the result equal to zero, and solving for m_k, obtains:

$$m_k = \frac{\mathrm{cov}(y, x_k)}{\mathrm{var}(x_k)} - \sum_{j \neq k} m_j \frac{\mathrm{cov}(x_j, x_k)}{\mathrm{var}(x_k)}. \qquad \text{(B4.13)}$$

This expression defines as many equations as there are slope parameters to be fitted, and these can be simultaneously solved to obtain values for every j. For example, if there are two predictors of y (say, x_1 and x_2) then the equations for the two slope parameters are:

$$m_1 = \frac{\mathrm{cov}(y, x_1)}{\mathrm{var}(x_1)} - m_2 \frac{\mathrm{cov}(x_1, x_2)}{\mathrm{var}(x_1)}, \quad \text{and} \qquad \text{(B4.14)}$$

$$m_2 = \frac{\mathrm{cov}(y, x_2)}{\mathrm{var}(x_2)} - m_1 \frac{\mathrm{cov}(x_1, x_2)}{\mathrm{var}(x_2)}. \qquad \text{(B4.15)}$$

Simultaneously solving (B4.14) and (B4.15) for m_1 and m_2 yields:

$$m_1 = \frac{\beta_{y,x_1} - \beta_{y,x_2}\beta_{x_2,x_1}}{1 - \rho_{12}^2}, \quad \text{and} \qquad \text{(B4.16)}$$

$$m_2 = \frac{\beta_{y,x_2} - \beta_{y,x_1}\beta_{x_1,x_2}}{1 - \rho_{12}^2}, \qquad \text{(B4.17)}$$

where $\rho_{12} = \mathrm{cov}(x_1,x_2)/\sqrt{(\mathrm{var}(x_1)\mathrm{var}(x_2))}$ is the correlation coefficient. In the main text, we use partial regression notation such that $m_1 = \beta_{y,x1|x2}$ and $m_2 = \beta_{y,x2|x1}$. Equations (B4.16) and (B4.17) show how these may be computed for any model directly from covariances, variances and correlation coefficients, i.e. without having to use least-squares minimization methods.

the cooperation phenotype, with a proportion p of individuals carrying an allele A which encodes the cooperator strategy, and the remaining $1 - p$ carrying an allele a which encodes the noncooperator strategy. Individuals are paired with a genetically identical social partner with probability α, and are paired with a random social partner with probability $1 - \alpha$. Social interaction mediates fecundity F, which is given by a unit baseline plus the individual's payoff from the game – payoffs are given in Table 3. Cooperators incur a fecundity cost C and provide their partners with a fecundity benefit B. Additionally, when two cooperators are paired, they both receive a synergistic benefit of D. We assume that the individual's relative fitness is equal to their relative fecundity: $v = F/\mathrm{E}(F)$.

Natural selection is formally defined according to change in genetic values (Box 1), and so we begin our analysis of this model by defining the genetic values of each type of individual (Box 2). We denote the phenotypic values of cooperators and noncooperators by $z = 1$ and $z = 0$ respectively (and hence $\mathrm{E}(z) = p$). The genic values can be chosen arbitrarily, so long as they are different. We choose $x = 1$ for the A allele and $x = 0$ for the a allele (and hence $\mathrm{E}(x) = p$). From equation (B2.1), we fit the linear model $\mathrm{E}(z|x) = \mathrm{E}(z) + \beta_{z,x}\,(x - \mathrm{E}(x))$ to the population data (summarized in Table 4) by the method of least squares, to obtain $\beta_{z,x} = 1$. From equation (B2.2), an individual's genetic value is given by $g = \mathrm{E}(z) + \beta_{z,x}\,(x - \mathrm{E}(x))$, and this is $g = 1$ for cooperators and $g = 0$ for noncooperators.

Having obtained genetic values, we can write a linear model of social-partner genetic value g' as a function of an individual's own genetic value g, to compute the coefficient of relatedness between social partners. From equation (B4.1), this is $\mathrm{E}(g'|g) = \mathrm{E}(g') + \beta_{g',g}\,(g - \mathrm{E}(g))$. Fitting this linear model to the population data (Table 5) obtains $\beta_{g',g} = \alpha$. Hence, the kin-selection coefficient of relatedness is given by:

Table 3 Fecundity payoffs in a two-player game of synergistic cooperation. Based upon the model of Queller (1985).

	Fecundity payoff, P	
	Cooperator	Noncooperator
Cooperator	$B - C + D$	$-C$
Noncooperator	B	0

Table 4 Population data for genetic value calculations in Queller's (1985) and van Veelen's (2009) models.

Frequency	Genic value (x)	Phenotype (z)
$1 - p$	0	0
p	1	1
	$\mathrm{E}(x) = p$	$\mathrm{E}(z) = p$

Table 5 Data for cost, benefit and relatedness calculations in Queller's (1985) model.

Frequency	Own genetic value (g)	Partner's genetic value (g')	Relative fitness (v)
$(1-p)^2 + \alpha p(1-p)$	0	0	$1/\mathrm{E}(F)$
$(1-\alpha)p(1-p)$	0	1	$(1+B)/\mathrm{E}(F)$
$(1-\alpha)p(1-p)$	1	0	$(1-C)/\mathrm{E}(F)$
$p^2 + \alpha p(1-p)$	1	1	$(1+B-C+D)/\mathrm{E}(F)$
	$\mathrm{E}(g) = p$	$\mathrm{E}(g') = p$	$\mathrm{E}(v) = 1$

$$r = \alpha. \tag{7}$$

Finally, we may write a linear model of the individual's relative fitness v as a function of her genetic value g and the genetic value of her social partner g'. From equation (4), this is $\mathrm{E}(v|g,g') = \mathrm{E}(v) + \beta_{v,g|g'}\ (g - \mathrm{E}(g)) + \beta_{v,g'|g}\ (g' - \mathrm{E}(g'))$. Fitting this linear model to the population data (Table 5) obtains the cost ($c = -\beta_{v,g|g'}$) and benefit ($b = \beta_{v,g'|g}$) terms of Hamilton's rule:

$$c = \tilde{C} - \frac{1}{1+\alpha}(\alpha + (1-\alpha)p)\tilde{D}, \quad \text{and} \tag{8}$$

$$b = \tilde{B} + \frac{1}{1+\alpha}(\alpha + (1-\alpha)p)\tilde{D} \tag{9}$$

where $\tilde{B} = B/\mathrm{E}(F)$, $\tilde{C} = C/\mathrm{E}(F)$ and $\tilde{D} = D/\mathrm{E}(F)$. Note that, of the synergistic relative fecundity benefit $\tilde{D}$ that accrues to the proportion $\alpha + (1-\alpha)p$ of cooperators who have cooperative social partners, a fraction $1/(1+\alpha)$ is assigned by the regression analysis to direct fitness (i.e. $-c = -\tilde{C} + [1/(1+\alpha)][\alpha + (1-\alpha)p]\ \tilde{D}$) and the remaining fraction $\alpha/(1+\alpha)$ is assigned to indirect fitness (i.e. $rb = \alpha\tilde{B} + [\alpha/(1+\alpha)][\alpha + (1-\alpha)p]\ \tilde{D}$).

Having computed the three terms (r, b and c) of Hamilton's rule, we can write a condition for natural selection to increase the average genetic value for cooperation. This is $rb - c > 0$, or:

$$\alpha B - C + (\alpha + (1-\alpha)p)D > 0. \tag{10}$$

Condition (10) is precisely that which was derived by Queller (1985) for cooperation to be favoured by natural selection. Although Queller (1985) interpreted this result as a demonstration of the failure of Hamilton's rule to correctly predict genetic change when gene action is not additive, we have derived it directly from Hamilton's rule. How are we to account for this difference in conclusion?

Queller (1985) identified the first two terms on the LHS of condition (10) (i.e. $\alpha B - C$, or equivalently, $rB - C$) as the LHS of Hamilton's rule. The condition $rB - C > 0$ does not correctly predict genetic change when there is a deviation from additivity ($D \neq 0$), so Queller (1985) suggested that deviations from additivity result in Hamilton's rule's failure. However, Hamilton's rule is $rb - c > 0$, i.e. the costs and benefits are fitness effects, properly defined by a least-squares multi-regression analysis, rather than arbitrary fecundity parameters that just happened to be represented by similar letters of the alphabet (Grafen, 1985b; Queller, 1992a; Rousset, 2004; Gardner *et al.*, 2007). We have shown that this canonical Hamilton's rule correctly predicts the action of natural selection, irrespective of the degree of synergy.

A general point here is that we are not suggesting that alternative partitions, such as that provided by Queller (1985), are incorrect or not useful. Indeed, Queller's result helps to make the mechanistic components of fitness effects more explicit (Gardner *et al.*, 2007). Instead, our point is that Hamilton's rule remains a correct statement, albeit one in which the cost and benefit terms may be somewhat complicated.

Three-player games

van Veelen (2009) has argued that, although the issue of synergy can be resolved for two-player games, Hamilton's rule cannot be guaranteed to give a correct prediction of the action of natural selection in games of three or more players. This is based upon the idea that Hamilton's rule contains fewer parameters than are needed to model such complexity. van Veelen (2009) illustrated this point with an example of a three-player 'stag hunt' game. Here, we provide a detailed reanalysis of this game, from a Hamilton's-rule perspective.

We consider an infinite population of haploid individuals, with cooperation strategies genetically encoded as in the two-player game, described above. We now assume that social groups contain three individuals, with the proportion of groups containing one, two and three cooperators being denoted p_1, p_2 and p_3 respectively (hence, the proportion of groups containing zero cooperators is $p_0 = 1 - p_1 - p_2 - p_3$, and the population frequency of cooperators is $p = 1/3\ p_1 + 2/3\ p_2 + p_3$. Again, we assume that social interaction mediates individual fecundity, with individuals having a unit baseline fecundity plus a payoff that is determined by their own strategy and the strategies of their two social partners, according to Table 6. According to the three-player stag-hunt game, assumed by van Veelen (2009), an individual loses a unit of fecundity if she cooperates, and additionally gains two units of fecundity if all three individuals in her social group cooperate. As before, we assume that an individual's relative fitness is given by $v = F/\mathrm{E}(F)$.

As we have assumed a genotype-phenotype map that is exactly the same as for the above analysis of Queller's (1985) model (Table 4), we already know that the genetic value for cooperation is $g = 1$ for individuals carrying the A allele and is $g = 0$ for individuals carrying the a allele. We compute relatedness by writing a linear model of social-partner genetic value g' as a function of an individual's own genetic value g. From equa-

Box 5 – Relatedness

The appropriate regression coefficient must be very near to Sewall Wright's Coefficient of Relationship – Hamilton (1963)

In the main text, we have defined the kin-selection coefficient of genetic relatedness between social partners as $r \equiv \mathrm{cov}(g',g)/\mathrm{cov}(g,g)$, where g is an individual's genetic value for a character of interest, g' is the genetic value of her social partner and the covariances are taken over all individuals in the population. This definition emerges directly from the application of Price's (1970) theorem to social evolution, as shown in equation (4) of the main text (see also Hamilton, 1970). The quantity $\mathrm{cov}(g',g)/\mathrm{cov}(g,g)$ has a simple interpretation within the context of the least-squares linear regression approach: it is equal to the least-squares linear regression coefficient $\beta_{g',g}$ in the model:

$$\mathrm{E}(g'|g) = \mathrm{E}(g') + \beta_{g',g}(g - \mathrm{E}(g)). \qquad \text{(B5.1)}$$

Hence, we can view the kin-selection coefficient of relatedness as arising from a least-squares linear regression analysis, in which we are attempting to predict the genetic value of an individual's social partner given that we know the individual's own genetic value for a character of interest.

Relatedness has been conceived as a regression coefficient from the very beginning of Hamilton's work on kin-selection theory. In the very first of these articles, Hamilton (1963) realized that the impact of a social partner's increased reproductive success, as a consequence of being the recipient of the actor's altruistic behaviour, on the genetic evolution of a population, does not simply depend upon the expected number of genes identical to those of the actor that will be added to the gene pool. In addition, those genes that are not identical to those of the actor and that will also be added to the gene pool will have a 'diluting' effect, reducing the extent to which the actor's genes increase in frequency within the population. Consequently, Hamilton (1963) pointed out that the appropriate coefficient of relatedness is one that measures the concentration of the actor's genes in the recipient, and he suggested that this would in principle take the form of a regression coefficient – albeit one that is well-approximated by Wright's correlation coefficient of relationship under certain simplifying assumptions (see also Hamilton, 1964, 1970). Orlove (1975) appears to have been the first to explicitly express the coefficient of relatedness as the slope of the least-squares linear regression line through data describing the genetics of actors and their recipients. Orlove & Wood (1978) were the first to explicitly express relatedness as the ratio of covariances of genetic values, reproduced above. The geometrical aspects of the regression approach to relatedness was developed by Grafen (1985a,b), and it forms the basis of the widely used Kinship software that provides estimates of relatedness from microsatellite data (Queller & Goodnight, 1989).

Orlove & Wood (1978) also provided an alternative conception of relatedness in terms of covariances of genetic and phenotypic values, which gained some traction in the literature (e.g. Michod & Hamilton, 1980; Seger, 1981; Michod, 1982). Queller (1992a) provided both purely genetical and also phenotypic covariance forms for relatedness, suggesting that Hamilton's rule is exact and completely general if the genetical form is used but special assumptions are required in order for the phenotypic form to provide the correct answer. Frank (1997a,b) showed that the kin-selection approach can be extended to give a general theory of 'correlated selection', in which any set of predictors that correlate with genes and fitness can be incorporated into the regression analysis, giving rise to relatedness-like coefficients that describe the association between gene and arbitrary predictor. However, throughout this article, we have focused upon purely genetical predictors, in the spirit of Fisher's (1918, 1930) paradigm. This is so that the associated partition of natural selection into direct and indirect components will exactly correspond with the direct and indirect components of inclusive fitness, the quantity that organisms are designed to maximize (Hamilton, 1964, 1996; Grafen, 2006a), which better clarifies the link between the process and purpose of Darwinian adaptation.

tion (B4.1), this is $\mathrm{E}(g'|g) = \mathrm{E}(g') + \beta_{g',g}\ (g - \mathrm{E}(g))$, and fitting this model to the population data (Table 7) obtains $\beta_{g',g} = 1 - (p_1 + p_2)/3p(1-p)$. Hence, the kin-selection coefficient of relatedness is:

$$r = 1 - \frac{p_1 + p_2}{3p(1-p)}. \qquad (11)$$

Finally, we may write a linear model of the individual's relative fitness v as a function of her genetic value g and the genetic values of her social partners g' and g''. This is $\mathrm{E}(v|g,g',g'') = \mathrm{E}(v) + \beta_{v,g|g',g''}\ (g - \mathrm{E}(g)) + \beta_{v,g'|g,g''}\ (g' - \mathrm{E}(g')) + \beta_{v,g''|g,g'}\ (g'' - \mathrm{E}(g''))$. Fitting this linear model to the population data (Table 7) obtains the cost ($c = -\beta_{v,g|g',g''}$) and benefit ($b = \beta_{v,g'|g,g''} + \beta_{v,g''|g,g'}$) terms of Hamilton's rule:

$$c = \left[1 - \frac{3p(1-p)}{9p(1-p) - 2(p_1 + p_2)} \times \frac{p_3}{p} \times 2\right] / \mathrm{E}(F), \quad \text{and} \qquad (12)$$

$$b = \left[\frac{6p(1-p)}{9p(1-p) - 2(p_1 + p_2)} \times \frac{p_3}{p} \times 2\right] / \mathrm{E}(F). \qquad (13)$$

Table 6 Fecundity payoffs in a three-player game of synergistic cooperation. Based upon the model of van Veelen (2009).

	Fecundity payoff, P		
	Non-C/Non-C	C/Non-C	C/C
Cooperator	−1	−1	1
Noncooperator	0	0	0

Table 7 Population data for cost, benefit and relatedness calculations in van Veelen's (2009) model.

Frequency	Own genetic value (g)	Partner's genetic value (g')	Other partner's genetic value (g'')	Relative fitness (v)
p_0	0	0	0	$1/\mathrm{E}(F)$
$p_1/3$	0	0	1	$1/\mathrm{E}(F)$
$p_1/3$	0	1	0	$1/\mathrm{E}(F)$
$p_2/3$	0	1	1	$1/\mathrm{E}(F)$
$p_1/3$	1	0	0	0
$p_2/3$	1	0	1	0
$p_2/3$	1	1	0	0
p_3	1	1	1	$2/\mathrm{E}(F)$
	$\mathrm{E}(g) = p$	$\mathrm{E}(g') = p$	$\mathrm{E}(g'') = p$	$\mathrm{E}(v) = 1$

Note that, of the synergistic relative fecundity benefit $2/\mathrm{E}(F)$ that accrues to the proportion p_3/p of cooperators who have two cooperative social partners, a fraction $1/(1 + 2r) = 3p(1 - p)/[9p(1 - p) - 2(p_1 + p_2)]$ is assigned by the regression analysis to direct fitness, and the remaining fraction $2r/(1 + 2r) = 6pr(1 - p)/[9p(1 - p) - 2(p_1 + p_2)]$ is assigned to indirect fitness, analogous to the result obtained in our reanalysis of Queller's (1985) model.

Having computed the three terms (r, b and c) of Hamilton's rule, we can now write a condition for natural selection to increase the average genetic value for cooperation. This is $rb - c > 0$, and simplifies to give the condition:

$$\frac{p_3}{p} > \frac{1}{2}, \tag{14}$$

Hence, if more cooperators are present in groups of three cooperators than are not, then natural selection favours an increase in cooperation.

This is exactly the result derived by van Veelen (2009). However, van Veelen (2009) claimed that this result cannot be formulated in terms of Hamilton's rule (in particular, 'cannot be given in a formula with costs, benefits and relatedness only'). He argued that the same is true of any model with three or more social partners interacting in a synergistic way. This conclusion follows from an argument that: (i) in such models, details of population structure in addition to relatedness (genetic similarity of pairs of individuals) must be known to predict the action of natural selection; and (ii) such details do not feature as terms in Hamilton's rule, so this cannot generally provide correct predictions. For example, two model populations can be set up such that the relatedness between pairs of social partners is the same but other moments of population structure are different, such that cooperation is selected for in one model but selected against in the other. van Veelen (2009) asserted that the corresponding Hamilton's rules for both models will be identical, and hence unable to account for the difference in selective outcome.

However, as we have shown, this is based upon an incorrect understanding of kin selection and Hamilton's rule. We have derived van Veelen's (2009) condition for cooperation to be selectively favoured, using Hamilton's rule. Irrespective of higher moments of population structure, Hamilton's rule correctly predicts the action of natural selection. This is because the cost and benefit terms are also mediated by population structure, so that although higher moments of genetic structure are not captured in the relatedness coefficient, they may exert their effects through the cost and benefit terms of Hamilton's rule. van Veelen (2009) has therefore confused the cost and benefit terms in Hamilton's rule with arbitrary payoff parameters in an evolutionary game (another recent example of this occurs in Nowak *et al.*, 2010b). In reality, the cost and benefit terms in Hamilton's rule are the direct and indirect average fitness effects of a genetic predisposition for a character of interest, which are prone to be mediated by the genetic and ecological milieu.

Altruism among nonkin

The previous section considered claims for the invalidity of Hamilton's rule that arise from misconceptions regarding costs and benefits. In this section, we consider a recent claim for the invalidity of Hamilton's rule that concerns the coefficient of genetic relatedness. Fletcher & Doebeli (2009) devised a thought experiment, which they claimed has shown that altruism can be favoured by natural selection even when relatedness is zero. Thus, Fletcher & Doebeli (2009) argue that relatedness is not the key to altruism, and Hamilton's rule can fail to correctly predict the action of natural selection. Here, we show that their conclusion arises because they have incorrectly computed the coefficient of relatedness in their model.

Fletcher & Doebeli (2009) considered a hypothetical strain of bacterium, for which an altruistic trait is encoded by two bi-allelic loci. Individuals with allele A at the first locus and allele b at the second locus, or allele a at the first locus and allele B at the second locus, have a probability q of committing suicide, an altruistic act that provides a social partner with a fecundity benefit B (see Table 8). Individuals with the a allele at the first locus and the b allele at the second locus never commit suicide,

Table 8 Fecundity payoffs in a game of altruistic suicide. Based upon the model of Fletcher & Doebeli (2009).

	Fecundity payoff, P	
	Suicider	Nonsuicider
Suicider	−1	−1
Nonsuicider	B	0

and individuals with the A allele at the first locus and the B allele at the second locus are presumed absent (owing to lethality of this genotype). Fletcher & Doebeli (2009) considered that an experimenter always pairs Ab individuals with aB individuals, and ab with ab (hence, there is an implicit constraint that the frequency of A and B alleles must be equal; we denote these frequencies by $p/2$, i.e. the frequency of individuals with either the Ab or the aB genotype is p). Fletcher & Doebeli (2009) showed that the genes for altruism (A and B) can be favoured, despite the recipients of altruism being dissimilar at both loci (Ab helps aB, or aB helps Ab) and hence, they claim, having a relatedness of zero.

As the preceding sections have made clear, natural selection is defined according to genetic – not phenotypic or genotypic – change. So, a kin selection analysis of this model begins by computing individuals' genetic values for the altruistic character in question. We denote: the altruistic suicide phenotype $z = 1$ and the nonsuiciding phenotype $z = 0$; the A allele $x_1 = 1$ and the a allele $x_1 = 0$; and the B allele $x_2 = 1$ and the b allele $x_2 = 0$. From equation (B2.1), we can write a linear regression model of the phenotype as a function of its genic predictors: $\mathrm{E}(z|x_1,x_2) = \mathrm{E}(z) + \beta_{z,x1|x2}\,(x_1 - \mathrm{E}(x_1)) + \beta_{z,x2|x1}\,(x_2 - \mathrm{E}(x_2))$ and, fitting this linear model to the population data (Table 9), we find the least-square regression coefficients are $\beta_{z,x1} = \beta_{z,x2} = q$. From equation (B2.2), genetic value is defined as $g = \mathrm{E}(z) + \beta_{z,x1}\,(x_1 - \mathrm{E}(x_1)) + \beta_{z,x2}\,(x_2 - \mathrm{E}(x_2))$, which is $g = 0$ for ab individuals ($x_1 = x_2 = 0$) and $g = q$ for both Ab individuals ($x_1 = 1$, $x_2 = 0$) and aB individuals ($x_1 = 0$, $x_2 = 1$).

Table 9 Population data for genetic value calculations in Fletcher & Doebeli's (2009) model.

Proportion	Genic value for locus 1 (x_1)	Genic value for locus 2 (x_2)	Phenotype (z)
$1 - p$	0	0	0
$p(1 - q)/2$	0	1	0
$pq/2$	0	1	1
$p(1 - q)/2$	1	0	0
$pq/2$	1	0	1
	$\mathrm{E}(x_1) = p/2$	$\mathrm{E}(x_2) = p/2$	$\mathrm{E}(z) = pq$

Having computed genetic values, we may compute the kin-selection coefficient of genetic relatedness. From equation (B4.1), we write a linear-regression model $\mathrm{E}(g'|g) = \mathrm{E}(g') + \beta_{g',g}(g - \mathrm{E}(g))$. Fitting this to population data (Table 10) obtains $\beta_{g',g} = 1$. Hence, the kin-selection coefficient of relatedness is given by:

$$r = 1. \tag{15}$$

We may now compute the cost and benefit terms of Hamilton's rule. From equation (4), we write a linear-regression model of relative fitness as a function of genetic predictors: $\mathrm{E}(v|g,g') = \mathrm{E}(v) + \beta_{v,g|g'}(g - \mathrm{E}(g)) + \beta_{v,g'|g}\,(g' - \mathrm{E}(g'))$ and, fitting this to the population data (Table 10), we obtain the cost ($c = -\beta_{v,g|g'}$) and benefit ($b = \beta_{v,g'|g}$) terms of Hamilton's rule:

$$c = \left[1 + \frac{1}{2}qB\right]/\mathrm{E}(F), \quad \text{and} \tag{16}$$

$$b = \left[B - \frac{1}{2}qB\right]/\mathrm{E}(F). \tag{17}$$

Note that, of the antagonistic relative-fecundity cost $qB/\mathrm{E}(F)$ that accrues to potential suiciders paired with other potential suiciders (owing to suciders being unable to enjoy the fecundity benefits of their partner's suicide), a fraction $1/(1+r) = ½$ is assigned by the regression analysis to direct fitness, and the remaining fraction $r/(1+r) = ½$ is assigned to indirect fitness, analogous to

Table 10 Population data for cost, benefit and relatedness calculations in Fletcher & Doebeli's (2009) model. Here we have generalized the degree of assortment $0 \le \alpha \le 1$; Fletcher's & Doebeli's model is recovered by setting $\alpha = 1$.

Frequency	Own genetic value (g)	Partner genetic value (g')	Own phenotype (z)	Partner phenotype (z')	Relative fitness (v)
$(1 - p)^2 + \alpha p(1 - p)$	0	0	0	0	$1/\mathrm{E}(F)$
$(1 - \alpha)p(1 - p)(1 - q)$	0	q	0	0	$1/\mathrm{E}(F)$
$(1 - \alpha)p(1 - p)q$	0	q	0	1	$(1+B)/\mathrm{E}(F)$
$(1 - \alpha)p(1 - p)(1 - q)$	q	0	0	0	$1/\mathrm{E}(F)$
$(1 - \alpha)p(1 - p)q$	q	0	1	0	0
$(p^2 + \alpha p(1 - p))(1 - q)^2$	q	q	0	0	$1/\mathrm{E}(F)$
$(p^2 + \alpha p(1 - p))q(1 - q)$	q	q	0	1	$(1+B)/\mathrm{E}(F)$
$(p^2 + \alpha p(1 - p))q(1 - q)$	q	q	1	0	0
$(p^2 + \alpha p(1 - p))q^2$	q	q	1	1	0
	$\mathrm{E}(g) = pq$	$\mathrm{E}(g') = pq$	$\mathrm{E}(z) = pq$	$\mathrm{E}(z') = pq$	$\mathrm{E}(v) = 1$

the results obtained in our reanalysis of Queller's (1985) and van Veelen's (2009) models.

Having calculated r, b and c, we can substitute these into Hamilton's rule $rb - c > 0$ to arrive at a condition for natural selection to favour an increase in altruistic suicide:

$$(1 - q)B > 1. \quad (18)$$

Thus, we have recovered Fletcher & Doebeli's (2009) result that natural selection can act to favour altruism in the context of their model. Fletcher & Doebeli (2009) argued that this result was not obtainable from a Hamilton's-rule perspective. However, we have derived their result directly from a Hamilton's-rule reanalysis of their model, and indeed Hamilton (1971) already provided an analysis of such 'assortation' effects.

The difference in conclusion arises because Fletcher & Doebeli (2009) incorrectly calculated the coefficient of genetic relatedness, treating it as a measure of genotypic similarity *per se*, rather than a measure of genetic similarity. Although social partners under their unusual pairing scheme carry different alleles at both of the loci underlying the altruistic trait, their genetic predisposition for altruism is exactly the same, and hence the coefficient of genetic relatedness is unity ($r = 1$), and not zero ($r = 0$) as Fletcher & Doebeli (2009) suggested. Hence, relatedness remains the key to altruism. More generally, Fletcher & Doebeli's (2009) thought that experiment can be seen as an example of when cooperation is favoured due to assortative interactions between altruists, as discussed in the greenbeard literature (reviewed by Gardner & West, 2010).

Although the formal genetical definition of relatedness used here may appear somewhat alien, it is in fact natural for a number of reasons. First, as students of organismal adaptation, we are not interested in molecular evolution *per se*, so it is understandable that descriptions of natural selection should not hinge upon molecular similarity. Indeed, Darwinism was developed and formalized before the molecular genetics revolution of the 1950s, so it should be no surprise that such details are extraneous to the basic theory (Grafen, 1999, 2009). Second, natural selection can only 'see' phenotypes to the extent that they are heritable (Fisher, 1930). Consequently, it is understandable that descriptions of natural selection should hinge upon the heritable portion of the phenotype (i.e. the individual's genetic value). Thirdly, relatedness is a population-level statistic, and should not change according to which individuals we choose to focus upon. Fletcher & Doebeli (2009) focused upon the pairing of genotypically dissimilar Ab and aB individuals, and neglected the concomitant pairing of genotypically identical ab individuals. If relatedness were measured genotypically, then the relatedness defined by this pairing scheme would lurch from $r = 0$ to $r = 1$, depending upon whose genotypes we focused our attention. Defining relatedness properly (i.e. genetically) obtains $r = 1$ for both Ab/aB and ab/ab pairings.

Discussion

Hamilton's (1963, 1964, 1970) partitions natural selection into two components: direct-fitness effects and indirect-fitness effects. Being a simple statement of natural selection, that makes no mechanistic assumptions of how these direct and indirect fitness effects come about, Hamilton's rule has the same generality as the theory of natural selection itself. In particular, Hamilton's rule does not break down in response to factors such as gene interactions, strong selection, frequency dependence, pure strategies, mechanisms that favour altruism between nongenealogical relatives, facultative strategies, greenbeards, evolutionary bifurcations, indirect genetic effects, fitness consequences over multiple generations or group selection or non-pairwise interactions (Table 2). All of these factors are implicitly allowed for in Hamilton's rule.

General theory vs. streamlined methodology

Evolutionary theory falls into two broad categories: that which aims for generality, and that which aims for applicability to specific problems. These two strands are present in kin-selection theory, and our main focus in this article has been to emphasize the general theory of kin selection and the corresponding general formulation of Hamilton's rule. However, kin-selection theorists also make use of streamlined methodologies which – owing to simplifying assumptions – provide straightforward tools that are readily applied to specific problems to yield concrete, testable results. These streamlined methodologies often employ the phenotypic gambit (Grafen, 1984; see below), i.e. the assumption of a straightforward genetic architecture, and vanishingly low genetic variation is often assumed so that regression coefficients can be replaced with derivatives (e.g. Taylor, 1996; Taylor & Frank, 1996; Frank, 1997b, 1998; Rousset & Billiard 2000; Rousset, 2004; Taylor *et al.*, 2007a; Box 6), allowing the modeller to draw upon the power of calculus.

We have employed Fisher's (1918, 1930) statistical, genetical paradigm to show that Hamilton's rule provides a completely general partition of natural selection. Although they have clear meanings, the cost, benefit and relatedness terms of Hamilton's rule are potentially very complicated functions of genotypic, demographic and ecological parameters (Grafen, 1985b, 2006a; Frank, 1998; Rousset, 2004; Gardner *et al.*, 2007). The cost and benefit terms are strictly defined as partial additive fitness effects obtained by a least-squares regression analysis, which fits straight lines through population data of arbitrary complexity (equations 5 & 6; Queller, 1992a; Frank, 1997a, 1998; Gardner *et al.*, 2007), and should not be confused with arbitrary model parameters such as fecundity effects or abstract utility payoffs in an evolutionary game (Grafen, 1985b; Rousset, 2004). Likewise, the coefficient of relatedness is a statistical term, emerging from least-squares linear regression analysis of social

Box 6 – How to make a kin-selection model

The standard maximization method for solving ESS problems can be adapted to analyse models with interactions among relatives – Taylor & Frank (1996, p. 28)

Optimality models of phenotypic evolution rest upon a level of abstraction that neglects the precise genetic architecture of a trait to facilitate the generation of testable predictions when such genetic details are unknown. This has been termed the 'phenotypic gambit' (Grafen, 1984), and it has been a hugely successful approach to developing evolutionary theory (Krebs & Davies, 1978, 1984, 1991, 1993, 1997).

From equation (2), the change in average genetic value owing to the action of natural selection is given by:

$$\Delta_S \mathrm{E}(g) = \beta_{v,g}\mathrm{var}(g), \tag{B6.1}$$

i.e. it is proportional to the least-squares linear regression of relative fitness against genetic value for the trait in question, where the coefficient of proportionality is the additive genetic variance. Assuming a differentiable functional relationship between phenotypic value for a trait z and relative fitness v (i.e. $v = W(z)$), and the simplest relationship between genetic and phenotypic values (i.e. $g = z$), then, in the limit of vanishingly small variation around the population average trait value $\bar{z}$, the regression is equal to the tangent of the fitness function, giving:

$$\Delta_S \mathrm{E}(g) = \left.\frac{\mathrm{d}W}{\mathrm{d}z}\right|_{z=\bar{z}} \mathrm{var}(g). \tag{B6.2}$$

The optimality approach typically begins by assuming (or deriving) a fitness function W that relates an individual's phenotypic trait value to its relative fitness, and differentiates this to obtain the direction of natural selection operating on this trait. If the tangent to the fitness function at the point $\bar{z}$ occupied by a population is positive ($\mathrm{d}W/\mathrm{d}z > 0$) then natural selection acts to favour the trait of interest, and if the tangent is negative ($\mathrm{d}W/\mathrm{d}z < 0$) then natural selection acts to disfavour the trait of interest. If an intermediate trait value is favoured by natural selection, then this is found by setting the derivative equal to zero and solving for $\bar{z}$ (and checking, from the second derivative, that this corresponds to a fitness maximum and not a fitness minimum). Hence, simple models relating trait value to fitness can be readily analysed using the tools of calculus to obtain predictions for evolutionarily and convergence stable strategies (see Taylor, 1996 for more details).

This basic approach has been adapted by Taylor & Frank (1996) for application to social evolutionary problems (see also Taylor, 1996; Frank, 1997b, 1998; Rousset, 2004; Taylor *et al.*, 2007a,b). They allow for the individual's fitness to be mediated by the phenotypes expressed by neighbours as well as by its own phenotype i.e. $v = W(z,z')$. Applying the chain rule of differential calculus, they partition the fitness derivative of equation (B6.2) into partial derivative effects of own and social partner phenotype:

$$\frac{\mathrm{d}W}{\mathrm{d}z} = \left.\frac{\partial W}{\partial z}\right|_{z=z'=\bar{z}} + \left.\frac{\partial W}{\partial z'}\right|_{z=z'=\bar{z}} \times \frac{\mathrm{d}z'}{\mathrm{d}z}. \tag{B6.3}$$

As above, in the limit of vanishing trait variation the regression terms in the personal fitness formulation of Hamilton's rule (equation 5) are equal to the derivative terms in equation (B6.3), i.e. $\beta_{v,g|g'} \to \partial W/\partial z|_{z=z'=\bar{z}} = -c$, $\beta_{v,g'|g} \to \partial W/\partial z'|z = z' = \bar{z} = b$ and $\beta_{g',g} \to d_{z'}/d_z = r$. Hence, the action of natural selection predicted by this simple chain rule approach corresponds exactly to Hamilton's rule: the trait is favoured when $rb - c > 0$, disfavoured when $rb - c < 0$ and intermediate optima satisfy $rb - c = 0$. Here, we have considered that all individuals in the model population are equivalent. However, class structure is readily allowed for by this simple approach (Taylor & Frank, 1996; and see Gardner *et al.*, 2009 for a recent worked example).

As an illustrative example, consider a simple model of the evolution of parasite virulence (Frank, 1996a). We assume a simple tradeoff between the exploitation of host resources for the production of parasite propagules vs. keeping the host alive to achieve sustained exploitation. We assume the overall transmission of infectious propagules from an infected host is a linearly decreasing function of the rate z' of exploitation by its parasites. And we assume that the share of these infectious propagules that derive from a focal parasite within the infection is given by its relative exploitation strategy within the host, z/z'. Hence, the focal parasite's transmission success is given by:

$$W(z,z') = \frac{z}{z'}(1 - az'). \tag{B6.4}$$

Applying the approach of Taylor & Frank (1996), we have partial derivatives $\partial W/\partial z|_{z=z'=\bar{z}} = (1 - a\bar{z})/\bar{z}$ and $\partial W/\partial z'|z = z' = \bar{z} = -1/\bar{z}$, and hence $\mathrm{d}W/\mathrm{d}z|_{z=z'=\bar{z}} = (1 - a\bar{z} - r)/\bar{z}$. Thus, the optimal level of exploitation satisfies $\mathrm{d}W/\mathrm{d}z|_{z=z'=\bar{z}} = 0$, i.e. $z^* = (1 - r)/a$. The reduction in host longevity owing to parasite exploitation is proportional to $V = az'$, and at the parasite optimum this is $V = 1 - r$. Thus, higher parasite relatedness r and higher host susceptibility to parasite exploitation a favour reduced exploitation by parasites. In this simple model, the level of host susceptibility a has no effect on the virulence suffered by the host at the parasite optimum: if hosts evolve reduced susceptibility to parasite exploitation, parasites respond by

Box 6 (continued)

increasing their level of exploitation until hosts are suffering as much as they were before (Frank, 1996a).

In this virulence example, we have assumed numerous parasites per infection, such that z' is both the average exploitation phenotype exhibited by the social partners of the focal parasite and also the average exploitation phenotype exhibited by the infection as a whole. However, for social groups comprising a smaller number of individuals, we must distinguish the average for social partners from the group average – as the latter contains the former and additionally includes a contribution from the focal individual herself. In particular, if the group contains the focal individual with phenotype z and $n - 1$ other individuals with average phenotype z', the group's average phenotypic value is $z'' = (1/n)z + ((n - 1)/n)z'$. For 'whole group' traits, whose social effects impact equally on all individual's the group, it is often easier to use z'' rather than z' as the second variable in the chain rule analysis. However, it is important to note that the corresponding relatedness coefficient dz''/dz is a whole-group relatedness, i.e. the relatedness of the focal individual to the average member of her group, including herself. In the absence of relatedness between social partners, whole-group relatedness will not be zero but rather $1/n$. In other scenarios, an 'others only' approach will be more natural. Importantly, the partition of natural selection made by Hamilton's rule (i.e. direct vs. indirect fitness effects) corresponds to the others-only partition (any impact of social group that owes to the individual being a part of its own group counts towards the direct, rather than the indirect, fitness effect) and the coefficient of relatedness in Hamilton's rule is an others-only relatedness (Frank, 1996b; Pepper, 2000).

The personal-fitness methodology developed by Taylor & Frank (1996) has hugely increased the ease of performing social-evolutionary analyses. Although an inclusive-fitness approach can be used to solve mathematical models, it can sometimes be tricky for the modeller to keep track of all the inclusive-fitness effects of an action, whereas it is often easier to write down a complete personal-fitness function and simply follow the recipe laid out by Taylor & Frank (1996). The ease with which the approach accommodates such complexities such as class structure and relatedness varying at multiple scales helps to ensure that biology leads the analysis. Moreover, the methodology allows analyses to be developed at the desired level of complexity – for example, the fitness function can be assumed or else derived from a mechanistic model, and relatedness can be treated as a parameter ('open' model) or else derived from an explicit model of population processes ('closed' model; Gardner & West, 2006).

partners' genetic predispositions for characters of interest (Hamilton, 1963, 1970, 1972; Orlove, 1975; Orlove & Wood, 1978; Michod & Hamilton, 1980; Grafen, 1985a; Queller, 1992a; Frank, 1997a, 1998; Box 5), and should not be confused with arbitrary quantities such as those relating to genealogical or genotypic or phenotypic closeness. All of these points follow if we regard Hamilton's rule as a straightforward extension of Fisher's (1918, 1930) genetical paradigm of natural selection, which represents the mathematical foundations of Darwinian theory itself. It is simply incorrect to claim that Hamilton's rule requires restrictive assumptions or that it almost never holds (e.g. van Veelen, 2009; Nowak *et al.*, 2010b).

By contrast, the use of more streamlined methodology is to develop testable predictions for specific problems. The general form of Hamilton's rule applies to any scatter of genetic and fitness data, irrespective of the underlying causes of this variation. It makes clear that, whatever the relationship between trait and reproductive success, the resulting action of natural selection can be decomposed into cost, benefit and relatedness terms. This is an important result to have obtained, but in most cases the aim of the evolutionary theorist will be to develop readily testable, concrete predictions for particular scenarios, and here the mechanistic causes of the relationship between trait and reproductive success will be of more interest. In such cases, it is more appropriate to build a mechanistic, dynamically sufficient model and to use any standard evolutionary methodology, such as neighbour-modulated-fitness, theoretical population genetics, quantitative genetics or evolutionary game theory to analyse the model and derive predictions (e.g. Taylor & Frank, 1996). Hamilton's rule may then be used as a conceptual aid for understanding the results delivered by these analyses, and connecting these with the results of other models analysed using different methods, as it provides a general framework and universal language for social-evolution theory (Taylor & Frank, 1996; Frank, 1998; Rousset, 2004; Gardner *et al.*, 2007). Thus, kin-selection theory is not a specific methodological approach for making and solving models, distinct from population genetics or evolutionary game theory (*contra* Nowak *et al.*, 2010b and Traulsen, 2010). Rather, it is the very foundation of social-evolution theory.

The streamlined methodologies used to analyse kin-selection models make simplifying assumptions, which are not required for Hamilton's rule to work but which

are made to achieve dynamical sufficiency or simply for ease of analysis. For example, selection distorts genealogies, such that relatedness can be difficult to compute in populations undergoing strong selection (Hamilton, 1964; Bulmer, 1994; but see Ajar, 2003; Gardner *et al.*, 2007 and Rousset & Roze, 2007), and so a standard simplifying assumption is that of weak selection, which enables neutral-population coefficients of relatedness to be used as close approximations to the real thing (Taylor 1992, 1996; Frank, 1986, 1998; Rousset, 2004). However, the frequent use of this assumption does not mean that relatedness is only defined in neutral populations or that Hamilton's rule fails when selection is not weak. More generally, very many authors have not been careful to separate the general theory of kin selection from the streamlined methods, when discussing the limiting assumptions of the latter. In Table 1, we have restricted ourselves to listing only those articles in which the suggested limitations of the kin-selection approach have concerned its foundations rather than the more heuristic versions of the theory.

Statistics and evolutionary theory

Our account of the general theory of kin selection has been framed in statistical terms. The use of statistical methods in the mathematical development of Darwinian theory has itself been subjected to recent criticism (van Veelen, 2005; Nowak *et al.*, 2010b), so we address this criticism here. The concern is that statistical terms – such as covariances and least-squares regressions – should properly be reserved for conventional statistical analyses, where hypotheses are tested against explicit data, and that they are out of place in the foundations of evolutionary theory (van Veelen, 2005; Nowak *et al.*, 2010b). However, this concern is misplaced. First, natural selection is a statistical process, and it is therefore natural that this should be defined in terms of aggregate statistics, even if only strictly by analogy (Frank, 1997a, 1998). Second, Fisher (1930, p198) coined the term 'covariance' in the context of his exposition of the genetical theory of natural selection, so the evolutionary usage of this term has precedent over the way the term is used in other fields. Third, it is a mistake to assume that it must be the evolutionary theorist, writing out covariances, who is performing the equivalent of a statistical analysis. A better analogy is to regard Mother Nature in the role of statistician, analysing fitness effects of genes by the method of least-squares, and driving genetic change according to the results of her analyses (cf. Crow, 2008). More generally, analogy is the basis of all understanding, so when isomorphisms arise unexpectedly between different branches of mathematics (in this case, theoretical population genetics and statistical least-squares analysis) this represents an opportunity for advancing scientific progress and not an anomaly that is to be avoided.

The phenotypic gambit

The phenotypic gambit is a simplifying assumption that phenotypes of biological interest are controlled by a simple genetic architecture, such that evolution proceeds on an essentially adaptationist basis without being upset by recombination between epistatically-interacting loci and other such factors (Grafen, 1984). This gambit is central to the streamlined kin-selection methodologies (Box 6), and the evolutionarily stable strategy (ESS) approach more generally. It is clearly a caricature of biological reality, and its users acknowledge that population genetics provides the formal basis for evolutionary theory, and that exact prediction of how a trait will evolve does require detailed knowledge of its true underlying genetic architecture. However, this concession is then set aside on the basis that such genetic details are not known for even a single trait in a single species, so that such exactness is unattainable in practice. Instead, the phenotypic gambit trades exactness for generality, and is employed to deliver approximate predictions that may be applied to a wide range of species that face similar adaptive problems but potentially encode their phenotypes with altogether different loci. Thus, the phenotypic gambit is a pragmatic approach to problems that evade exact solution.

However, and more importantly, the phenotypic gambit facilitates the interplay between theory and data. First, working out the exact solution to a specific model is not very helpful, because in reality we will have a cloud of possible models in mind, and want to get at their robust and general features. The phenotypic gambit provides a robust approximation to a wide range of models, that can tell us more than the exact solution of any one model (for a discussion of such issues more generally in mathematical biology, see May, 2004). Second, the phenotypic gambit provides readily derived predictions, the robustness of which can be easily investigated in response to variation in both the number and values of model parameters. For example, because population-genetic methods require specific equations for each particular set of assumptions about how relatives interact, the ESS approach has been the only general way to study sex allocation for arbitrary population structures (Frank, 1995, 1998). It is for this reason that kin-selection theory has had such a larger impact than population-genetic theory on empirical areas such as sex allocation (West, 2009, pp. 359–361). More generally, decades of theoretical and empirical research activity in disciplines such as behavioural and evolutionary ecology have demonstrated the clear usefulness of the phenotypic gambit (Krebs & Davies, 1978, 1984, 1991, 1997; Cockburn, 1991; Stearns, 1992; Alcock, 2005; Mayhew, 2006; Westneat & Fox, 2010).

We have taken a subtly different approach in this article, using Price's (1970, 1972b) theorem to explicitly separate the selection and transmission components of

evolutionary change, and setting aside the transmission effect and focusing upon the selection effect. Thus, we have not made any limiting assumptions about the genetic architecture underlying traits of interest, but we have paid for this generality by limiting ourselves to only partial results, corresponding to only the selective component of evolutionary change. Hamilton's rule is explicitly understood to be a theorem that concerns natural selection only, and not the totality of evolutionary change. Hamilton's rule would be trivially violated by factors such as mutation pressure if we were to understand it as a strong statement about evolutionary change as a whole. Given that Hamilton's rule is a statement about natural selection only, then it must be formally framed in terms of genetic change, rather than genotypic or phenotypic change, as natural selection only 'sees' the heritable component of traits. The phenotypic gambit provides a means of directly translating these genetical results into phenotypic terms that are of more immediate use to empirical biologists.

Kin selection

The term 'kin selection' is used in at least three different ways. Hamilton (1963, 1964, 1970, 1971, 1975) developed Hamilton's rule as a general partition of natural selection, as we have described above, with r being genetic relatedness. The term 'kin selection' was introduced in a short commentary by Maynard Smith (1964), to distinguish Hamilton's theory from Wynne-Edwards (1962) ideas about group-level adaptations. Since then, kin selection has been used in three ways: (i) traits are described as being favoured by kin selection if they are favoured due to the consequences of interactions between close genealogical relatives (i.e. the rb term is important, and r is defined in terms of genealogy; e.g. Maynard Smith, 1964; Grafen, 2006a); (ii) traits are described as being favoured by kin selection if they are favoured due to the indirect-fitness consequences of the trait (i.e. the rb term is important, and r is defined as relative genetic similarity *per se*, following Hamilton, 1970; e.g. Foster *et al.*, 2006); (iii) kin-selection theory or the kin-selection partition describes how the total action of natural selection can be separated into two components – direct and indirect selection (e.g. Frank, 1998).

The difference between the first two uses is that the second (and third) allows for when indirect benefits can accrue owing to interaction between individuals who are not genealogically close relatives, such as through greenbeards effects (Hamilton, 1964; Dawkins, 1976; Gardner & West, 2010). Although the first use fits best with the use of 'kin' in the English language, it is the second use that best describes Hamilton's theory. Although this could be used to argue that kin selection is not the best phrase to describe Hamilton's theory (see also Hamilton, 1975), its usage is far too common for dropping it to be a realistic option. By way of analogy, the term 'linkage disequilibrium' refers to any statistical association between genes, and not just those that are due to physical linkage of genes on the same chromosome. However, rather than get rid of the term, population geneticists credit each other with enough intelligence to see past the misnomer, and do not waste time trying to bring down hitchhiking theory by pointing out examples where genes are situated on different chromosomes. Furthermore, in practice, the distinction between relatedness and relationship is rarely important, because the major cause of significant relatedness between individuals in the real world is genealogical closeness (Okasha, 2002), with greenbeard effects providing only very infrequent exceptions (reviewed by Gardner & West, 2010). Nonetheless, confusion is best avoided by realizing that different authors may mean different things by the phrase kin selection (or, indeed, Hamilton's rule), and by being clear about what is intended by the phrase, whenever it is used (West *et al.*, 2007a,b).

Alternative partitions

Other ways of partitioning natural selection are possible, and indeed it can often be useful to look at the same process from different perspectives (Maynard Smith, 1983; Frank, 1998). Here, we have focused upon the kin-selection partition of natural selection into direct and indirect components. An alternative partition is that of multilevel selection, which partitions total change into between-group and within-group components (Price, 1972b; Hamilton, 1975; Wade, 1985; Okasha, 2006). Another is the Barton-Turelli multilocus methodology, which partitions natural selection into a sum of products of selection coefficients and generalized linkage disequilibria (Barton & Turelli, 1991; Kirkpatrick *et al.*, 2002; Gardner *et al.*, 2007).

The usefulness of these different partitions depends upon the purpose at hand. The major use of the kin-selection approach is that it facilitates the linking of theory to empirical data, as demonstrated by research in fields such as sex allocation, cooperation and conflict in animal societies, cooperation in microbes, parent-offspring conflict, sibling conflict, selfish genetic elements and genomic imprinting (Mock & Parker, 1997; Haig, 2002; Sachs *et al.*, 2004; Burt & Trivers, 2006; Ratnieks *et al.*, 2006; West *et al.*, 2006, 2007a,b; Strassmann & Queller, 2007; Boomsma, 2009; Clutton Brock, 2009; Hatchwell, 2009; West, 2009). The reason for this empirical success is that the kin-selection approach emphasises how an individual's behaviour influences her own reproductive success (and hence her direct fitness, c) and also the reproductive success of her social partners (and hence her indirect fitness, rb). Thus, we can imagine the individual adjusting her inclusive fitness (the sum of her direct and indirect fitness) by

altering her behaviour. Natural selection leads organisms to appear designed to maximize their inclusive fitness (Hamilton, 1964, 1996; Grafen, 2006a; Gardner, 2009) and, as a consequence, kin-selection theory provides an excellent means for understanding the rationale of social adaptations.

Acknowledgments

We thank: N. Barton, J. Biernaskie, S. Frank, A. Grafen, A. Herre, L. Lehmann, S. Lion, N. Mideo, A. Moore, H. Ohtsuki, I. Pen, D. Queller, A. Rodrigues and F. Rousset for helpful discussion; the Royal Society, Balliol College, Leverhulme Trust and ERC for funding.

References

Ajar, E. 2003. Analysis of disruptive selection in subdivided populations. *BMC Evol. Biol.* **3**: 22.

Alcock, J. 2005. *Animal Behavior*, 8th edn. Sinauer Associates, Sunderland, MA.

Alizon, S. & Taylor, P.D. 2008. Empty sites can promote altruistic behaviour. *Evolution* **62**: 1335–1344.

Avilés, L. 1993. Interdemic selection and the sex ratio: a social spider perspective. *Am. Nat.* **142**: 320–345.

Barton, N.H. & Turelli, M. 1991. Natural and sexual selection on many loci. *Genetics* **127**: 229–255.

Bijma, P. & Wade, M.J. 2008. The joint effects of kin, multilevel selection and indirect genetic effects on response to genetic selection. *J. Evol. Biol.* **21**: 1175–1188.

Boomsma, J.J. 2009. Lifetime monogamy and the evolution of eusociality. *Philos. Trans. R. Soc. B Biol. Sci.* **364**: 3191–3207.

Bourke, A.F.G. & Franks, N.R. 1995. *Social Evolution in Ants*. Princeton University Press, Princeton, NJ.

Bulmer, M.G. 1994. *Theoretical Evolutionary Ecology*. Sinauer Associates, Sunderland, MA.

Burt, A. & Trivers, R. 2006. *Genes in Conflict*. Harvard University Press, Cambridge, MA.

Cavalli-Sforza, L.L. & Feldman, M.W. 1978. Darwinian selection and ''altruism''. *Theor. Popul. Biol.* **14**: 268–280.

Charlesworth, B. 1980. Models of kin selection. In: *Evolution of Social Behaviour: Hypotheses and Empirical Tests* (H. Markl, ed), pp. 11–26. Verlag, Weinheim.

Charlesworth, B. 1994. *Evolution in Age-Structured Populations*. Cambridge University Press, Cambridge.

Charnov, E.L. 1977. An elementary treatment of the genetical theory of kin selection. *J. Theor. Biol.* **66**: 541–550.

Charnov, E.L. 1978. Evolution of eusocial behavior: offspring choice or parental parasitism? *J. Theor. Biol.* **75**: 451–465.

Charnov, E.L. 1982. *The Theory of Sex Allocation*. Princeton University Press, Princeton, NJ.

Cheverud, J.M. 1984. Evolution by kin selection: a quantitative genetic model illustrated by maternal performance in mice. *Evolution* **38**: 766–777.

Clutton Brock, T. 2009. Structure and function in mammalian societies. *Philos. Trans. R. Soc. B Biol. Sci.* **364**: 3229–3242.

Cockburn, A. 1991. *An Introduction to Evolutionary Ecology*. Wiley-Blackwell, Oxford.

Colwell, R.K. 1981. Group selection is implicated in the evolution of female-biased sex ratios. *Nature* **290**: 401–404.

Craig, R. 1979. Parental manipulation, kin selection, and the evolution of altruism. *Evolution* **33**: 319–334.

Crow, J.F. 2008. Commentary: Haldane and beanbag genetics. *Int. J. Epidemiol.* **37**: 442–445.

Crozier, R.H. 1992. The genetic evolution of flexible strategies. *Am. Nat.* **139**: 218–223.

Darwin, C.R. 1859. *On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life*. John Murray, London.

Dawkins, R. 1976. *The Selfish Gene*. Oxford University Press, Oxford.

Dawkins, R. 1982. *The Extended Phenotype*. Oxford University Press, Oxford.

Doebeli, M. & Hauert, C. 2006. Limits to Hamilton's rule. *J. Evol. Biol.* **19**: 1386–1388.

Edwards, A.W.F. 1994. The fundamental theorem of natural selection. *Biol. Rev. Camb. Philos. Soc.* **69**: 443–474.

Falconer, D.S. 1981. *Introduction to Quantitative Genetics*, 2nd edn. Longmans Green, London.

Fisher, R.A. 1918. The correlation between relatives on the supposition of mendelian inheritance. *Trans. R. Soc. Edinb.* **52**: 399–433.

Fisher, R.A. 1930. *The Genetical Theory of Natural Selection*. Clarendon Press, Oxford.

Fisher, R.A. 1941. Average excess and average effect of a gene substitution. *Ann. Eugen.* **11**: 53–63.

Fletcher, J.A. & Doebeli, M. 2006. How altruism evolves: assortment and synergy. *J. Evol. Biol.* **19**: 1389–1393.

Fletcher, J.A. & Doebeli, M. 2009. A simple and general explanation for the evolution of altruism. *Proc. R. Soc. Lond. B Biol. Sci.* **276**: 13–19.

Fletcher, J.A. & Zwick, M. 2006. Unifying the theories of inclusive fitness and reciprocal altruism. *Am. Nat.* **168**: 252–262.

Fletcher, J.A., Zwick, M., Doebeli, M. & Wilson, D.S. 2006. What's wrong with inclusive fitness? *Trends Ecol. Evol.* **21**: 597–598.

Foster, K.R., Wenseleers, T. & Ratnieks, F.L.W. 2006. Kin selection is the key to altruism. *Trends Ecol. Evol.* **21**: 57–60.

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

Frank, S.A. 1995. George Price's contributions to evolutionary genetics. *J. Theor. Biol.* **175**: 373–388.

Frank, S.A. 1996a. Models of parasite virulence. *Q. Rev. Biol.* **71**: 37–78.

Frank, S.A. 1996b. Policing and group cohesion when resources vary. *Anim. Behav.* **52**: 1163–1169.

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

Frank, S.A. 1997b. Multivariate analysis of correlated selection and kin selection, with an ESS maximization method. *J. Theor. Biol.* **189**: 307–316.

Frank, S.A. 1998. *Foundations of Social Evolution*. Princeton University Press, Princeton, NJ.

Gardner, A. 2009. Adaptation as organism design. *Biol. Lett.* **5**: 861–864.

Gardner, A. & Grafen, A. 2009. Capturing the superorganism: a formal theory of group adaptation. *J. Evol. Biol.* **22**: 659–671.

Gardner, A. & West, S.A. 2006. Demography, altruism and the benefits of budding. *J. Evol. Biol.* **19**: 1707–1716.

Gardner, A. & West, S.A. 2010. Greenbeards. *Evolution* **64**: 25–38.

Gardner, A., West, S.A. & Barton, N.H. 2007. The relation between multilocus population genetics and social evolution theory. *Am. Nat.* **169**: 207–226.

Gardner, A., Arce, A. & Alpedrinha, J. 2009. Budding dispersal and the sex ratio. *J. Evol. Biol.* **22**: 1036–1045.

Gintis, H. 2000. Strong reciprocity and human sociality. *J. Theor. Biol.* **206**: 169–179.

Goodnight, C., Rauch, E., Sayama, H., de Aguiar, M. A. M., Baranger, M. & Bar-Yam, Y. 2008. 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**: 23–44.

Grafen, A. 1984. Natural selection, kin selection and group selection. In: *Behavioural Ecology: An Evolutionary Approach*, 2nd edn (J.R. Krebs & N.B. Davies, eds), pp. 62–84. Blackwell, Oxford.

Grafen, A. 1985a. A geometric view of relatedness. *Oxford Surveys in Evolutionary Biology* **2**: 28–90.

Grafen, A. 1985b. Hamilton's rule OK. *Nature* **318**: 310–311.

Grafen, A. 1999. Formal Darwinism, the individual-as-maximising-agent analogy, and bet-hedging. *Proc. R. Soc. Lond. B* **266**: 799–803.

Grafen, A. 2000. Developments of the Price equation and natural selection under uncertainty. *Proc. R. Soc. Lond. B Biol. Sci.* **267**: 1223–1227.

Grafen, A. 2002. A first formal link between the Price equation and an optimization program. *J. Theor. Biol.* **217**: 75–91.

Grafen, A. 2006a. Optimization of inclusive fitness. *J. Theor. Biol.* **238**: 541–563.

Grafen, A. 2006b. A theory of Fisher's reproductive value. *J. Math. Biol.* **53**: 15–60.

Grafen, A. 2007a. Detecting kin selection at work using inclusive fitness. *Proc. R. Soc. Lond. B* **274**: 713–719.

Grafen, A. 2007b. An inclusive fitness analysis of altruism on a cyclical network. *J. Evol. Biol.* **20**: 2278–2283.

Grafen, A. 2009. Formalizing Darwinism and inclusive fitness theory. *Philos. Trans. R. Soc. B* **364**: 3135–3141.

Haig, D. 2002. *Genomic Imprinting and Kinship*. Rutgers University Press, New Brunswick, NJ.

Hamilton, W.D. 1963. The evolution of altruistic behavior. *Am. Nat.* **97**: 354–356.

Hamilton, W.D. 1964. The genetical evolution of social behaviour I & II. *J. Theor. Biol.* **7**: 1–52.

Hamilton, W.D. 1967. Extraordinary sex ratios. *Science* **156**: 477–488.

Hamilton, W.D. 1970. Selfish and spiteful behaviour in an evolutionary model. *Nature* **228**: 1218–1220.

Hamilton, W.D. 1971. Selection of selfish and altruistic behaviour in some extreme models. In: *Man and Beast: Comparative Social Behavior* (J. F. Eisenberg & W. S. Dillon, eds), pp. 57–91. Smithsonian Press, Washington, DC.

Hamilton, W.D. 1972. Altruism and related phenomena, mainly in social insects. *Ann. Rev. Ecol. Syst.* **3**: 193–232.

Hamilton, W.D. 1975. Innate social aptitudes of man: an approach from evolutionary genetics. In: *Biosocial Anthropology* (R. Fox, ed), pp. 133–155. Wiley, New York.

Hamilton, W.D. 1988. This week's citation classic. *Curr. Cont.* **40**: 16.

Hamilton, W.D. 1996. *Narrow Roads of Gene Land Volume 1 – Evolution of Social Behaviour*. W.H. Freeman, Oxford.

Hardy, I.C.W. 2002. *Sex ratios: Concepts and Research Methods*. Cambridge University Press, Cambridge.

Hatchwell, B.J. 2009. The evolution of cooperative breeding in birds: kinship, dispersal and life history. *Philos. Trans R. Soc. B Biol. Sci.* **364**: 3217–3227.

Killingback, T., Bieri, J. & Flatt, T. 2006. Evolution in group-structured populations can resolve the tragedy of the ccommons. *Proc. R. Soc. Lond. B* **273**: 1477–1481.

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

Krebs, J.R. & Davies, N.B. 1978. *Behavioural Ecology: An Evolutionary Approach*, 1st edn. Blackwell, Oxford.

Krebs, J.R. & Davies, N.B. 1984. *Behavioural Ecology: An Evolutionary Approach*, 2nd edn. Blackwell, Oxford.

Krebs, J.R. & Davies, N.B. 1991. *Behavioural Ecology: An Evolutionary Approach*, 3rd edn. Blackwell, Oxford.

Krebs, J.R. & Davies, N.B. 1993. *An Introduction to Behavioural Ecology*, 3rd edn. Blackwell, Oxford.

Krebs, J.R. & Davies, N.B. 1997. *Behavioural Ecology: An Evolutionary Approach*, 4th edn. Blackwell, Oxford.

Lehmann, L. 2010. Space-time relatedness and Hamilton's rule for long lasting behaviors in viscous populations. *Am. Nat.* **175**: 136–143.

Lehmann, L. & Keller, L. 2006a. The evolution of cooperation and altruism: a general framework and classification of models. *J. Evol. Biol.* **19**: 1365–1376.

Lehmann, L. & Keller, L. 2006b. Synergy, partner choice and frequency dependence: their integration into inclusive fitness theory and their interpretation in terms of direct and indirect fitness effects. *J. Evol. Biol.* **19**: 1426–1436.

Lehmann, L., Keller, L. & Sumpter, D. 2007a. The evolution of helping and harming on graphs: the return of the inclusive fitness effect. *J. Evol. Biol.* **20**: 2284–2295.

Lehmann, L., Keller, L., West, S.A. & Roze, D. 2007b. Group selection and kin selection: two concepts but one process. *Proc. Natl Acad. Sci. USA* **104**: 6736–6739.

Lynch, M. & Walsh, J.B. 1998. *Genetics and Analysis of Quantitative Traits*. Sinauer Associates, Sunderland, MA.

May, R.M. 2004. Uses and abuses of mathematics in biology. *Science* **303**: 790–793.

Mayhew, P. J. 2006. *Discovering Evolutionary Ecology*. Oxford University Press, Oxford.

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

Maynard Smith, J. 1980. Power and limits of optimization. In: *Evolution of Social Behaviour: Hypotheses and Empirical Tests* (H. Markl, ed), pp. 27–34. Verlag, Weinheim.

Maynard Smith, J. 1983. Models of evolution. *Proc. R. Soc. Lond. B Biol. Sci.* **219**: 315–325.

McGlothlin, J.W., Moore, A.J., Wolf, J.B. & Brodie, E.D. III 2010. Interacting phenotypes and the evolutionary process III: social evolution. *Evolution* **64**: 2558–2574.

Michod, R.E. 1982. The theory of kin selection. *Annu. Rev. Ecol. Syst.* **13**: 23–55.

Michod, R.E. & Hamilton, W.D. 1980. Coefficients of relatedness in socio-biology. *Nature* **288**: 694–697.

Mock, D.W. & Parker, G.A. 1997. *The Evolution of Sibling Rivalry*. Oxford University Press, Oxford.

Nowak, M.A. 2006. Five rules for the evolution of cooperation. *Science* **314**: 1560–1563.

Nowak, M.A., Tarnita, C.E. & Antal, T. 2010a. Evolutionary dynamics in structured populations. *Philos. Trans. R. Soc. Lond. B* **365**: 19–30.

Nowak, M.A., Tarnita, C.E. & Wilson, E.O. 2010b. The evolution of eusociality. *Nature* **466**: 1057–1062.
Okasha, S. 2002. Genetic relatedness and the evolution of altruism. *Philos. Sci.* **69**: 138–149.
Okasha, S. 2006. *Evolution and the Levels of Selection*. Oxford University Press, Oxford.
Orlove, M.J. 1975. A model of kin selection not invoking coefficients of relationship. *J. Theor. Biol.* **49**: 289–310.
Orlove, M.J. & Wood, C.L. 1978. Coefficients of relationship and coefficients of relatedness in kin selection: a covariance form for the RHO formula. *J. Theor. Biol.* **73**: 679–686.
Parker, G.A. 1989. Hamilton's rule and conditionality. *Ethol. Ecol. Evol.* **1**: 195–211.
Pepper, J.W. 2000. Relatedness in trait group models of social evolution. *J. Theor. Biol.* **206**: 355–368.
Price, G.R. 1970. Selection and covariance. *Nature* **227**: 520–521.
Price, G.R. 1972a. Fisher's 'fundamental theorem' made clear. *Ann. Hum. Genet.* **36**: 129–140.
Price, G.R. 1972b. Extension of covariance selection mathematics. *Ann. Hum. Genet.* **35**: 485–490.
Price, G.R. 1995. The nature of selection. *J. Theor. Biol.* **175**: 389–396.
Price, G.R. & Smith, C.A.B. 1972. Fisher's Malthusian parameter and reproductive value. *Ann. Hum. Genet.* **36**: 1–7.
Queller, D.C. 1984. Kin selection and frequency dependence: a game theoretic approach. *Biol. J. Linn. Soc.* **23**: 133–143.
Queller, D.C. 1985. Kinship, reciprocity, and synergism in the evolution of social behaviour. *Nature* **318**: 366–367.
Queller, D.C. 1992a. A general model for kin selection. *Evolution* **46**: 376–380.
Queller, D.C. 1992b. Quantitative genetics, inclusive fitness and group selection. *Am. Nat.* **139**: 540–558.
Queller, D.C. & Goodnight, K.F. 1989. Estimating relatedness using genetic markers. *Evolution* **34**: 258–275.
Queller, D.C. & Strassmann, J.E. 2006. Models of cooperation. *J. Evol. Biol.* **19**: 1410–1412.
Ratnieks, F.L.W., Foster, K.R. & Wenseleers, T. 2006. Conflict resolution in insect societies. *Annu. Rev. Entomol.* **51**: 581–608.
Rice, S.H. 2004. *Evolutionary Theory*. Sinauer Associates, Sunderland, MA.
Ridley, M. & Grafen, A. 1981. Are green beard genes outlaws? *Anim. Behav.* **29**: 954–955.
Robertson, A. 1966. A mathematical model of the culling process in dairy cattle. *Anim. Prod.* **8**: 95–108.
Robertson, A. 1968. The spectrum of genetic variation. In: *Population Biology and Evolution* (R.C. Lewontin, ed), pp. 5–16. Syracuse University Press, New York.
Ross-Gillespie, A., Gardner, A., West, S.A. & Griffin, A.S. 2007. Frequency dependence and cooperation: theory and a test with bacteria. *Am. Nat.* **170**: 331–342.
Rousset, F. 2004. *Genetic Structure and Selection in Subdivided Populations*. Princeton University Press, Princeton, NJ.
Rousset, F. & Billiard, S. 2000. A theoretical basis for measures of kin selection in subdivided populations: finite populations and localized dispersal. *J. Evol. Biol.* **13**: 814–825.
Rousset, F. & Ronce, O. 2004. Inclusive fitness for traits affecting metapopulation demography. *Theor. Popul. Biol.* **65**: 127–141.
Rousset, F. & Roze, D. 2007. Constraints on the origin and maintenance of genetic kin recognition. *Evolution* **61**: 2320–2330.
Sachs, J.L., Mueller, U.G., Wilcox, T.P. & Bull, J.J. 2004. The evolution of cooperation. *Q. Rev. Biol.* **79**: 135–160.
Seger, J. 1981. Kinship and covariance. *J. Theor. Biol.* **91**: 191–213.
Sheynin, O.B. 1993. On the history of the principle of least squares. *Arch. Hist. Exact Sci.* **46**: 39–54.
Stearns, S. C. 1992. *Evolution of Life Histories*. Oxford University Press, Oxford.
Strassmann, J.E. & Queller, D.C. 2007. Insect societies as divided organisms: the complexities of purpose and cross-purpose. *Proc. Natl Acad. Sci. USA* **104**: 8619–8626.
Taylor, P.D. 1988. Inclusive fitness models with two sexes. *Theor. Popul. Biol.* **34**: 145–168.
Taylor, P.D. 1989. Evolutionary stability in one-parameter models under weak selection. *Theor. Popul. Biol.* **36**: 125–143.
Taylor, P.D. 1990. Allele frequency change in a class-structured population. *Am. Nat.* **135**: 95–106.
Taylor, P.D. 1992. Altruism in viscous populations – an inclusive fitness model. *Evol. Ecol.* **6**: 352–356.
Taylor, P.D. 1996. Inclusive fitness arguments in genetic models of behaviour. *J. Math. Biol.* **34**: 654–674.
Taylor, P.D. & Frank, S.A. 1996. How to make a kin selection model. *J. Evol. Biol.* **180**: 27–37.
Taylor, P.D., Wild, G. & Gardner, A. 2007a. Direct fitness or inclusive fitness: how should we model kin selection? *J. Evol. Biol.* **20**: 301–309.
Taylor, P.D., Day, T. & Wild, G. 2007b. Evolution of cooperation in a finite homogeneous graph. *Nature* **447**: 469–472.
Traulsen, A. 2010. Mathematics of kin- and group-selection: formally equivalent? *Evolution* **64**: 316–323.
Traulsen, A. & Nowak, M.A. 2006. Evolution of cooperation by multilevel selection. *Proc. Natl Acad. Sci. USA* **103**: 10952–10955.
Traulsen, A., Shoresh, N. & Nowak, M.A. 2008. Analytical results for individual and group selection of any intensity. *Bull. Math. Biol.* **70**: 1410–1424.
Uyenoyama, M. & Feldman, M.W. 1980. Theories of kin and group selection: a population genetics perspective. *Theor. Popul. Biol.* **17**: 380–414.
Uyenoyama, M. & Feldman, M.W. 1981. On relatedness and adaptive topography in kin selection. *Theor. Popul. Biol.* **19**: 87–123.
Uyenoyama, M., Feldman, M.W. & Mueller, L.D. 1981. Population genetic theory of kin selection: multiple alleles at one locus. *Proc. Natl Acad. Sci. USA* **78**: 5036–5040.
van Veelen, M. 2005. On the use of the Price equation. *J. Theor. Biol.* **237**: 412–426.
van Veelen, M. 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.
Wade, M.J. 1985. Soft selection, hard selection, kin selection, and group selection. *Am. Nat.* **125**: 61–73.
Wenseleers, T. 2006. Modelling social evolution: the relative merits and limitations of a Hamilton's rule-based approach. *J. Evol. Biol.* **19**: 1419–1422.
West, S.A. 2009. *Sex Allocation*. Princeton University Press, Princeton, NJ.
West, S.A. & Gardner, A. 2010. Altruism, spite and greenbeards. *Science* **327**: 1341–1344.
West, S.A., Griffin, A.S., Gardner, A. & Diggle, S. P. 2006. Social evolution theory for microbes. *Nat. Rev. Microbiol.* **4**: 597–607.

West, S.A., Griffin, A.S. & Gardner, A. 2007a. Social semantics: altruism, cooperation, mutualism, strong reciprocity and group selection. *J. Evol. Biol.* **20**: 415–432.

West, S.A., Diggle, S.P., Buckling, A., Gardner, A. & Griffin, A.S. 2007b. The social lives of microbes. *Ann. Rev. Ecol. Evol. Syst.* **38**: 53–77.

West, S.A., Griffin, A.S. & Gardner, A. 2008. Social semantics: how useful has group selection been? *J. Evol. Biol.* **21**: 374–385.

West, S.A., El Mouden, C. & Gardner, A. 2011. Sixteen common misconceptions about the evolution of cooperation in humans. *Evol. Hum. Behav.*, doi: 10.1016/j.evolhumbehav.2010.08.001.

Westneat, D.F. & Fox, C.W. 2010. *Evolutionary Behavioral Ecology*. Oxford University Press, Oxford.

Wild, G., Gardner, A. & West, S.A. 2009. Adaptation and the evolution of parasite virulence in a connected world. *Nature* **459**: 983–986.

Wilson, E.O. 2005. Kin selection as the key to altruism: its rise and fall. *Soc. Res.* **72**: 159–166.

Wilson, D.S. 2008. Social semantics: towards a genuine pluralism in the study of social behaviour. *J. Evol. Biol.* **21**: 368–373.

Wilson, D.S. & Wilson, E.O. 2007. Rethinking the theoretical foundation of sociobiology. *Q. Rev. Biol.* **82**: 327–348.

Wolf, J.B. 2003. Genetic architecture and evolutionary constraint when the environment contains genes. *Proc. Natl Acad. Sci. USA* **100**: 4655–4660.

Wolf, J.B., Brodie, E.D. III, Cheverud, J.M., Moore, A.J. & Wade, M.J. 1998. Evolutionary consequences of indirect genetic effects. *Trends Ecol. Evol.* **13**: 64–69.

Wolf, J.B., Brodie, E.D. III & Moore, A.J. 1999. Interacting phenotypes and the evolutionary process II: selection resulting from social interactions. *Am. Nat.* **153**: 254–266.

Wynne-Edwards, V.C. 1962. *Animal Dispersion in Relation to Social Behaviour*. Oliver & Boyd, Edinburgh.

Received 22 November 2010; revised 5 January 2011; accepted 7 January 2011