Gene Mobility and the Concept of Relatedness

Jonathan Birch
Christ's College, University of Cambridge
jgb37@cam.ac.uk

(forthcoming in *Biology and Philosophy*; the final publication will be available at link.springer.com)

Abstract: Cooperation is rife in the microbial world, yet our best current theories of the evolution of cooperation were developed with multicellular animals in mind. Hamilton's theory of inclusive fitness is an important case in point: applying the theory in a microbial setting is far from straightforward, as social evolution in microbes has a number of distinctive features that the theory was never intended to capture. In this article, I focus on the conceptual challenges posed by the project of extending Hamilton's theory to accommodate the effects of gene mobility. I begin by outlining the basics of the theory of inclusive fitness, emphasizing the role that the concept of relatedness is intended to play. I then provide a brief history of this concept, showing how, over the past fifty years, it has departed from the intuitive notion of genealogical kinship to encompass a range of generalized measures of genetic similarity. I proceed to argue that gene mobility forces a further revision of the concept. The reason in short is that, when the genes implicated in producing social behaviour are mobile, we cannot talk of an organism's genotype *simpliciter*; we can talk only of an organism's genotype at a particular stage in its life cycle. We must therefore ask: with respect to which stage(s) in the life cycle should relatedness be evaluated? For instance: is it genetic similarity at the time of social interaction that matters to the evolution of social behaviour, or is it genetic similarity at the time of reproduction? I argue that, strictly speaking, it is neither of these: what really matters to the evolution of social behaviour is *diachronic* genetic similarity between the producers of fitness benefits at the time they produce them and the recipients of those benefits at the end of their life-cycle. I close by discussing the implications of this result. The main payoff is that it makes room for a possible new mechanism for the evolution of altruism in microbes that does not require correlated interaction among bearers of the genes for altruism. The importance of this mechanism in nature remains an open empirical question.

---

1 Introduction

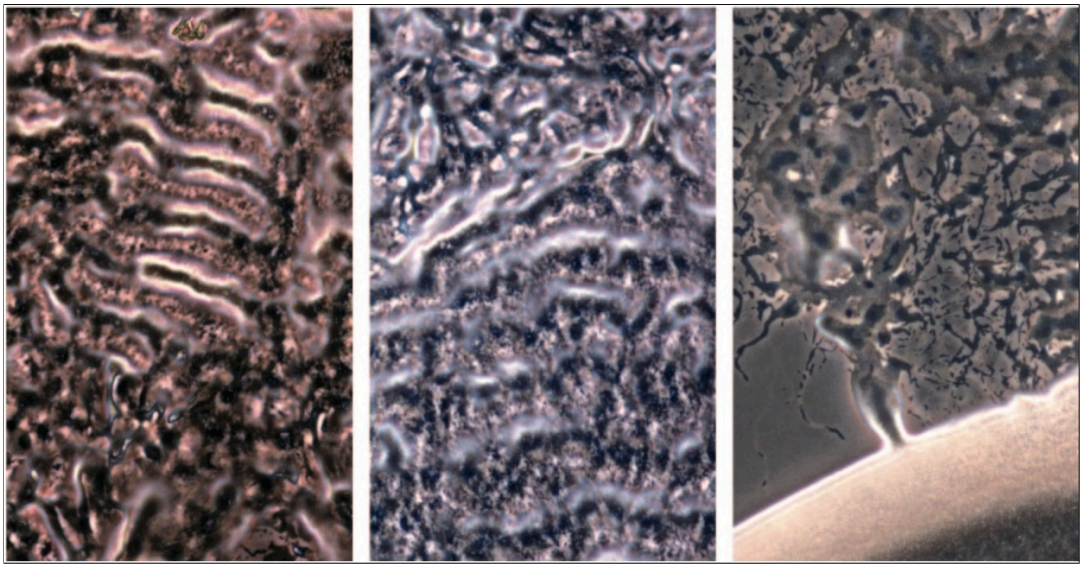
The most celebrated examples of cooperation in the natural world involve multicellular animals. We are awestruck (and rightly so) by the elaborate nests of ants, bees, wasps and termites; by the hunting strategies of dolphins, orcas and wolves; and by the complex social hierarchies of baboons and chimpanzees. Yet this stock of familiar examples provides only a meagre sample of the full range of cooperative phenomena in nature. Indeed, it is not just a small sample, but a biased one: many of nature's most spectacular social phenomena go largely unnoticed by human eyes, for they are too small for the unaided eye to see.

Consider, for example, the social amoeba *Dictyostelium discoideum* (Bonner 1959; Strassmann et al. 2000; Strassmann and Queller 2011). For much of their life cycle, these amoebae conform to our usual expectations of amoebae: they live in the soil, they engulf bacteria, they divide mitotically. When food gets scarce, however, things get interesting: if the amoebae are present in sufficient density, the starving amoebae aggregate to form a mobile 'slug'. The slug moves as one—and moves further and faster than any individual amoeba ever would—in the direction of heat and light. On reaching a favourable location, the slug stops and begins to transform into a fruiting body (Figure 1(a)). Around a fifth of the amoebae sacrifice their lives in this process, forming a hardy, cellulose stalk of dead cells. The remaining four fifths cluster at the tip of the stalk, where they generate and release spores. The spores are dispersed through the environment, reducing the probability that the amoebae they ultimately produce will encounter the same harsh conditions their parents endured.

For a second example, consider the social bacterium *Myxococcus xanthus* (Velicer and Vos 2009). *M. xanthus* too can generate fruiting bodies through the aggregation of free-living cells, but this is not the only trick up its sleeve. Recent work by James E. Berleman and colleagues (Berleman et al. 2006) has documented a mysterious behaviour in which the bacteria move collectively in a synchronized 'ripple' formation, like waves on the sea (Figure 1(b)). There is good evidence that this rippling is a predatory behaviour, triggered by the proximity of food; but the question remains open as to what predatory advantage, if any, it provides for the bacteria. One plausible hypothesis is that the formation is a kind of battle tactic: by rippling underneath a prey colony, the *M. xanthus* swarm is able to disrupt and dislodge its prey more effectively. If this is correct, then sophisticated 'pack hunting' techniques can no longer be considered the sole preserve of large, multicellular mammals.

Such impressive feats of collective action should convince us that cooperation in microbes is no mere sideshow: it is no less sophisticated, no less spectacular, and no less central to the life cycles of the organisms involved than cooperation in the macroscopic world. But these headline-grabbing examples are only the tip of the iceberg. The moral from the past few decades of research in microbiology is that sociality is pervasive in microbial populations (Crespi 2001; West et al. 2007a). We now realize that what looked like a blob on a Petri dish is in reality a dynamic social network:

(a) Fruiting bodies of the social amoeba *Dictyostelium discoideum* (photograph by Scott Solomon, reproduced with permission).



(b) Predatory ripple formation in *Myxococcus xanthus* (photographs from Berleman et al. 2006, reproduced with permission from the American Society for Microbiology).

Figure 1: Social microbes

a community in which vast numbers of microorganisms—often including members of several different species—interact with each other in complex and evolutionarily significant ways. Perhaps the most frequent type of social interaction involves the production and consumption of so-called 'public goods'. These goods include enzymes, surfactants, antibiotics, toxins, adhesive polymers, ammonia, and other useful organic molecules: what they have in common is that, when emitted by a producer into the external environment, they confer a fitness benefit on other nearby microorganisms. In recent years, hugely many instances of public goods production have been documented across otherwise highly disparate microbial taxa (West et al. 2007a).

Economic metaphors such as 'collective action' and 'public goods production' do more than simply point to a superficial resemblance between microbial sociality and our own feats of cooperation. They also draw our attention to a distinctive explanatory puzzle these behaviours present, a puzzle of a kind we also encounter in economics. For it is a truism, in both economics and evolutionary biology, that collective action and the production of public goods are often destabilized by free-riding (Olson 1965; Hardin 1982; Kagel and Roth 1995). Cooperative actions are often costly; free-riders reap the profits of their neighbours' cooperative actions without paying that cost; and so one intuitively expects such free-riders to do better, in the long run, than cooperators. There is no reason to think that microbial cooperation is immune to this threat. In microbes, the production of public goods usually imposes a metabolic cost (and hence, all else being equal, a fitness cost) on the producer; and so one would intuitively expect organisms that consume the goods without producing them to prosper at the producers' expense. So the puzzle we face is this: why does natural selection appear, in so many cases, to have favoured collective action and public goods production in microbes? Why are these behaviours not destabilized by free-riding? This is the 'problem of cooperation' in microbes (West et al. 2006, 2007a), and understanding how the problem is solved in nature is one of the central tasks of the emerging programme of 'sociomicrobiology' (Parsek and Greenberg 2005).

Of course, the problem of stabilizing cooperation under natural selection has been studied for over half a century in mainstream behavioural ecology, where theorists have developed a variety of modelling approaches designed to show how natural selection can, under certain conditions, favour cooperative behaviour (for reviews, see Gardner and Foster 2008; Wenseleers et al. 2010). Ideally, then, we would simply import our off-the-shelf solutions to the problem of cooperation to a microbial context. Matters, however, are not quite so simple. Complications arise from the fact that our traditional approaches to the problem were developed with multicellular animals in mind. Applying them to microbial populations is often far from straightforward, because social evolution in microbes has a number of distinctive features that our traditional theories were never intended to accommodate. As a result, we often find that our 'off-the-shelf' solutions require various extensions and amendments before they can be put to work in a microbial setting (Smith et al. 2010; Cornforth

et al. 2012).

W. D. Hamilton's (1964, 1970) theory of inclusive fitness is an important case in point. The details of the theory will be explained below; for now, it is enough to note that the theory assigns a crucial role to *relatedness* in the explanation of social behaviour. In populations of multicellular animals, the concept of relatedness is fairly (though not wholly) unproblematic: roughly speaking, the degree of relatedness between two organisms can be estimated by looking at the degree to which they have ancestors in common. As we will see below, this leaves room for disagreement about the best quantitative measure of relatedness, and in recent decades traditional pedigree-based measures have increasingly given way to generalized, statistical measures of genetic (and, in some cases, phenotypic) similarity. But as long as multicellular animals are the explanatory target, the differences between these measures can often be ignored in practice, and the intuitive notion of relatedness as genealogical kinship remains adequate for many purposes. This is far from the case in microbial populations. The source of all the trouble is the propensity of microbes to exchange genes horizontally, bypassing the parent-offspring channel. A growing body of evidence points to the importance of this process to social evolution in microbes (Smith 2001; West et al. 2007a; Nogueira et al. 2009; Rankin et al. 2011a, b; Mc Ginty et al. 2011; Mc Ginty et al. 2013; West and Gardner 2013). In this article, I will argue that even generalized statistical measures of relatedness do not fully account for the impact of gene mobility, forcing a further revision to the concept.

The outline of the article is as follows. In Section 2, I outline the basics of the theory of inclusive fitness, emphasizing the role that the concept of relatedness is intended to play. I then provide a brief introduction to this concept, showing how, over the past fifty years, it has departed from the intuitive notion of genealogical kinship to encompass a range of generalized measures of genetic similarity. In Sections 3, I explain why social evolution theorists cannot afford to ignore the phenomenon of gene mobility. In Sections 4-6, I proceed to argue that gene mobility forces a further revision of the concept of relatedness. The reason in short is that, when the genes implicated in producing social behaviour are mobile, we can no longer talk of an organism's genotype *simpliciter*; we can talk only of an organism's genotype *at a particular stage in its life cycle*. This raises the question: with respect to which stage in the life cycle should relatedness be defined? For instance: is it genetic similarity between organisms at the time of interaction that matters to the evolution of social behaviour, or is it genetic similarity at the time of reproduction? I use simple formal models to argue that, strictly speaking, it is neither: what really matters to the evolution of social behaviour is *diachronic* genetic similarity between the producers of fitness benefits at the time they produce them and the recipients of those benefits at the end of their life-cycle. In Section 7, I consider the wider significance of this result, highlighting the important empirical questions that a purely conceptual argument inevitably leaves open. In Section 8, I draw together the discussion and conclude.

As a final introductory remark, I should note that I am by no means the first author to stress

the radical consequences gene mobility holds for the conceptual foundations of evolutionary theory. Other authors have rightly emphasized the ways in which mobile genetic elements (particularly when they cross species boundaries) challenge traditional conceptions of the tree of life, biodiversity, individuality and other aspects of biological ontology (see, in particular, the articles collected in O'Malley 2010, Dupré 2012, and O'Malley 2013). This article can be seen as a contribution to this growing body of work. Nevertheless, the specific issue I address (i.e., the impact of gene mobility on the concept of relatedness) has not, to my knowledge, received any previous philosophical attention. My aim, then, is to start a new debate rather than to weigh into an existing one.

2 The Meaning and Measure of Relatedness

2.1 Relatedness as an exchange rate

As legend has it, the pithiest expression of the idea of inclusive fitness was made long before the term itself was coined. When asked if he would dive into a river to rescue a drowning stranger, the geneticist and co-architect of the modern synthesis J. B. S. Haldane is said to have replied: 'No, but I would do it for two brothers or eight cousins' (Maynard Smith 1964, 1976). If the story is true [1], then Haldane had latched on to a profound and powerful insight about social evolution: when interacting organisms share genes, the organisms may, in certain circumstances, have an evolutionary incentive to help one another. Even more remarkably, he grasped that the incentive comes in degrees, and that the size of the incentive is proportional to one's degree of relatedness to the potential beneficiary.

We owe the formal embodiment of this insight to W. D. Hamilton (1963, 1964, 1970, 1971, 1972, 1975). Today, Hamilton's theory lies at the heart of an established and burgeoning research programme, the explanatory domain of which has steadily expanded over recent decades (Bourke 2011a; Davies et al. 2012). The principle at the heart of the theory is 'Hamilton's rule', a deceptively simple statement of the conditions under which natural selection will favour the genes for a social behaviour. What it says, in a nutshell, is that the genes for a social behaviour are favoured by selection when $rB - C > 0$, where C is the fitness cost the behaviour imposes on the actor who performs it, B is the (average) fitness benefit the behaviour confers on other affected individuals (i.e. 'recipients'), and r is the (average) degree of relatedness between the recipients and the actor.

In Hamilton's rule, the concept of relatedness plays a distinctive theoretical role. Its role is that of an exchange rate, or conversion factor, which tells us how large the benefit conferred by a social behaviour on an actor's relatives has to be (in relation to the cost incurred by the actor) before the genes for that behaviour will be selected. Its role, in other words, is to quantify the *value* of a particular

[1] It may well be. The precise wording of the 'brothers and cousins' remark does not appear anywhere in Haldane's published work, but he does discuss broadly similar ideas in a 1955 article (Haldane 1955). Indeed, he explicitly considers a hypothetical scenario in which one must choose whether or not to jump into a river to save a drowning child.

recipient for a particular actor as an indirect means of transmitting the genes for a social behaviour to the next generation (Frank 1998). If the value of the recipient to the actor is large enough, then the genetic representation the actor gains indirectly by performing the behaviour in question may outweigh the genetic representation it loses by sacrificing a portion of its own reproductive success. In such cases, we can say that the behaviour contributes to the 'inclusive fitness' of the actor, even though it detracts from its direct fitness (Hamilton 1964; Grafen 2006; Bourke 2011a).

We can think of relatedness *qua* technical term in inclusive fitness theory as being implicitly defined by this theoretical role (cf. Lewis 1970). In inclusive fitness theory, relatedness just *is* that coefficient which provides the appropriate exchange rate for assessing the value to the actor (as an indirect means of transmitting the genes for a social behaviour to the next generation) of a fitness effect that falls on another individual. We can call this the 'exchange rate' conception of relatedness. Note, however, that this minimal definition leaves open the question of how this 'exchange rate' is to be formally defined and measured—and this is where complications start to arise.

2.2 Pedigree-based measures

There is, of course, an ordinary, intuitive concept of relatedness that pre-dates the theory of inclusive fitness. The intuitive notion of relatedness concerns the closeness of a genealogical relationship: I am more closely related to my sister than to my aunts, more closely related to my aunts than to my cousins, and so on. We can call this the 'pedigree' conception of relatedness. Long before Hamilton arrived on the scene, population geneticists had developed a variety of quantitative measures of pedigree relatedness for other theoretical purposes (Fisher 1918; Wright 1922; Li and Sacks 1954; Cockerham 1954; Kempthorne 1954, 1955, 1957; Falconer 1961; Haldane and Jayakar 1962; see Provine 1971 for historical detail).

Notable among these measures was Sewall Wright's (1922) 'coefficient of relationship', which quantifies the pedigree relatedness between two organisms on the basis of genealogical trees and background data about the extent of inbreeding. Suppose that we want to calculate Wright's coefficient of relationship (r_{AB}) between an organism A and another organism B. To apply Wright's method, we look at each of the genealogical paths linking A to B via a common ancestor.[2] To each path we assign a path coefficient, which depends on (i) the ploidy of the organisms, (ii) the number of parent-offspring connections on the path and (iii) the individual coefficients of inbreeding for A, B, and the common ancestor in question. To compute r_{AB}, we take the sum of the path coefficients over all our paths. In the special case of a population of diploid organisms with no inbreeding, the path coefficient for the i^{th} path reduces to 0.5^{n_i} , where n_i is the number of parent-offspring

[2] As Seger (1981) points out, we will find huge numbers of such paths between any two organisms if we go back far enough. There is thus an implicit pragmatic dimension to Wright's procedure: we look only at those genealogical relationships between organisms that are recent enough to have non-negligible path coefficients.

connections along that path. Hence, in this special case, $r = \sum_i 0.5^{n_i}$, and we obtain the standard pedigree measures of r: 0.5 between a parent and its offspring and between full siblings; 0.25 between grandparents and their grandchildren, between half siblings, and between aunts and uncles and their nephews and nieces; 0.125 between first cousins; and so on.

Both Haldane (in his 'brothers and cousins' quip) and Hamilton (in his original 1964 formulation of inclusive fitness theory) rely on a close connection between the 'pedigree' and 'exchange rate' conceptions of relatedness. Both, in fact, assume that Wright's coefficient of relationship (or something close to it) provides the right exchange rate for determining whether a social behaviour is in the evolutionary interest of the actor who performs it. In calculating whether it is worthwhile to jump into the river, Haldane implicitly assumes that $r = 1/2$ for siblings and $r = 1/8$ for first cousins, and these r-values line up with the standard coefficients of relationship for an outbred population of diploid organisms. Hamilton (1964), meanwhile, explicitly takes r to be 'equal to Sewall Wright's Coefficient of Relationship [...] unity for clonal individuals, one-half for sibs, one-quarter for half-sibs, one-eighth for cousins, ... and finally zero for all individuals whose relationship can be considered negligibly small' (Hamilton 1964, pp. 3-8).

When we are concerned with populations of multicellular animals, the assumption that the r-term in Hamilton's rule can be approximated by Wright's coefficient of relationship (or some other pedigree-based measure) is often reasonable. It is crucial, however, to appreciate the contingency of the connection between the 'pedigree' and 'exchange rate' conceptions of relatedness. There can be no *a priori* guarantee that the value (as a route to genetic representation in the next generation) of a particular recipient to a particular actor is reliably estimated by the degree of genealogical kinship between the two organisms; so there can be no *a priori* guarantee that traditional, pedigree-based measures of relatedness give us the appropriate exchange rate to use in an inclusive fitness calculation. Indeed, it is now widely recognized (and was recognized by Hamilton from the beginning) that there are possible cases in which the two conceptions come apart (Hamilton 1963, 1964, 1970, 1972, 1975).

Richard Dawkins's (1976) famous 'greenbeard' thought experiment (based on a remark in Hamilton 1964) provides one example. Dawkins asks us to consider a gene or gene-complex that causes its bearers to (a) grow a green beard, (b) recognize other bearers on the basis of their green beards, and (c) differentially help these individuals at a cost to themselves. Could such a trait ever be favoured by natural selection? Dawkins argues that it could. The key consideration is that, by causing its bearers to help other bearers of the same allele, the greenbeard gene may indirectly increase its genetic representation in the next generation, in spite of the cost it imposes on its bearer. In inclusive fitness terms, this argument relies on the idea that there is positive relatedness—in the 'exchange rate' sense—between social partners with respect to the greenbeard locus. Note, however, that the mechanism does not require any *pedigree* relatedness between bearers of the greenbeard gene: we can

suppose, if we want to, that the initial bearers of the trait are genealogically unrelated, and that the greenbeard mutant appears independently in each of them. The example is hypothetical, of course; but strikingly similar effects (mediated not by literal green beards, but by phenotypic markers or other associative mechanisms playing a similar role) have since been discovered empirically (Gardner and West 2010; West and Gardner 2010). Dawkins himself suggested that greenbeard effects would be rare in nature, as selection would favour genes at other loci that suppressed the expression of a greenbeard gene. This turns out to be true under some but not all conditions (Ridley and Grafen 1981; Gardner and West 2010; West and Gardner 2010; Biernaskie et al. 2011), and has evidently not been enough to prevent greenbeard effects evolving in some cases.

One might reasonably question, however, whether greenbeard effects are common enough to threaten the continuing utility of pedigree-based measures of r. Arguably a more serious problem for pedigree-based measures is that they are appropriate only on the assumption of weak selection (Hamilton 1963, 1964, 1970, 1972, 1975; Michod and Hamilton 1980; Uyenoyama and Feldman 1981; Toro et al. 1982; Grafen 1985). The need for this assumption arises because selection distorts the symmetry of family trees, making individuals with certain allele combinations more likely to survive (and contribute to future generations) than those with other combinations. The stronger selection gets, the more severe this distortion effect becomes. Suppose, for instance, that there is very strong selection against some allele G, and suppose that our focal (adult) organism happens to possess G. Now consider: in what fraction of this organism's surviving full siblings should we expect to find a copy of G? Intuitively, we want to say 0.5; but if selection against G is so strong that very few individuals with G survive to adulthood, the true fraction is likely to be far lower. Hence, when selection is strong, pedigree-based measures of relatedness fail to track the actual patterns of genetic similarity in the population. Consequently, they fail to track the true value of a particular recipient to a particular actor as an indirect route to genetic representation in the next generation. Here too, then, the 'pedigree' and 'exchange rate' conceptions of relatedness part ways.

2.3 Generalized statistical measures

These concerns about pedigree-based measures of r led a number of inclusive fitness theorists (including Hamilton) to develop generalized, statistical measures of the resemblance between social partners (Hamilton 1971, 1972, 1975; Orlove 1975; Orlove and Wood 1978; Michod and Hamilton 1980; Seger 1981; Grafen 1985; Queller 1985). These generalized measures, based on the statistical concepts of covariance and regression, tell us the correct exchange rate for the calculation of inclusive fitness (and hence the right predictions about the evolution of social behaviour) across a wider range of cases, including those cases in which (due to strong selection or greenbeard-type phenomena) pedigree-based measures give the wrong answers.

The intuitive thought behind these generalized measures is that we can visualize relatedness as the slope of the regression line—or 'line of best fit'—through a set of population data about individuals and their social partners. The most common measure of r in contemporary inclusive fitness theory regresses recipient genotypes on actor genotypes (Wenseleers et al. 2010; Gardner et al. 2011). More formally, we equate r with the β term in the following linear regression equation, where α represents the intercept of the regression line, g_i represents the additive genetic value of the i^{th} individual with respect to the social character of interest, $\hat{g}_i$ represents the average additive genetic value of the i^{th} individual's social partners, and ϵ represents the deviation from the regression line:

$$g_i = \alpha + \beta \hat{g}_i + \epsilon \tag{1}$$

The notion of an 'additive genetic value' is introduced here as a quantitative characterization of an individual's genotype. Roughly speaking, an individual's additive genetic value with respect to a character z is a measure of its genetic predisposition to express z. More precisely, it is its value for that character as predicted by a linear combination of its alleles, weighted by their average effects (sensu Fisher 1930, 1941) on the character (for further detail, see Falconer and Mackay 1996; Frank 1998 and Gardner et al. 2011).

Because equation (1) is a simple (one-predictor) regression, the regression coefficient β can be straightforwardly and usefully rewritten as the covariance (across all individuals) between g_i and $\hat{g}_i$, divided by the variance in g_i:

$$r = \beta = \frac{\text{Cov}\left(g_i, \hat{g}_i\right)}{\text{Var}\left(g_i\right)} \tag{2}$$

Hamilton's (1963, 1970) early work on inclusive fitness clearly states that r should be conceptualized as a regression coefficient, but he only spelt out the regression definition in detail in later work (1971, 1972). Although we usually cannot measure the relevant statistical quantities directly, we can estimate them in various ways. In many cases (especially in studies of social insects and social microbes) we can estimate them using molecular markers (Queller and Goodnight 1989). We thereby obviate the need for genealogical trees, bypassing the difficulties that arise in estimating r from considerations of pedigree.

In fact, equation (2) represents just one of a family of generalized statistical measures of relatedness in contemporary social evolution theory, and the best measure to use depends on the details of the case. I do not intend to catalogue all the options here, but I do need to highlight one particularly important variation. This alternative measure quantifies r in terms of the relationship between an individual's genes and the genetic composition of the *whole social group* in which that individual is situated (Breden 1990; Frank 1997, 2006; Pepper 2000). Formally, we equate relatedness with the

proportion of the variance in g_i that is accounted for by variance in the group means, G_i:

$$r_G = \frac{\text{Var}\,(G_i)}{\text{Var}\,(g_i)} \tag{3}$$

The above variance ratio is commonly known as the 'whole-group relatedness' (Pepper 2000). This measure comes into its own in models of public goods production in group-structured populations, because the amount of the public good received by any particular individual typically depends on the amount produced by the whole group (including that individual's own contribution) rather than on the amount produced by the individual's social partners alone (Pepper 2000; Frank 2006). One may be surprised to find mention of group-level properties (albeit averages, rather than 'emergent' group characters) in the foundations of inclusive fitness theory. This is indicative of the degree to which the inclusive fitness and multi-level approaches to the analysis of social evolution have converged in recent decades. Indeed, the formal affinities between the two approaches are now so strong that it is hard to locate any substantive differences between them (Wenseleers et al. 2010; Marshall 2011; though cf. West and Gardner 2013).

These generalized, statistical measures of r share a number of technical features that distinguish them from traditional pedigree-based measures. Firstly, they are generally *character-specific*: they quantify not the genetic similarity between social partners at *every* locus in the genome, but rather the genetic similarity at those loci which are relevant to the phenotypic character of interest. This character-specificity is implicitly smuggled into definitions (2) and (3) by the notion of an additive genetic value. This is because an additive genetic value, being a measure of an organism's genetic predisposition to express a particular phenotypic character, can only be evaluated once the relevant character is specified. This point can be legitimately overlooked when (as in most populations of multicellular animals) the genetic similarity between social partners is the same whichever loci we consider. But it cannot be overlooked when greenbeard effects are at work, and it cannot be overlooked in a microbiological context, since (as we will see) gene mobility can lead to highly locus-specific patterns of genetic similarity.

Secondly, these measures are *population-relative*: that is, they quantify the *differential* similarity between an individual and its social partner, relative to the population average. One consequence of this is that if everyone in the population is genetically identical at a particular locus, then the relatedness in the population with respect to that locus is formally undefined. In this respect, formal measures of relatedness are akin to formal measures of heritability.

Thirdly, covariances and regression coefficients can be negative as well as positive, so defining r in these terms allows for the possibility of *negative relatedness*. In what sort of biological scenario might a negative value of relatedness actually arise? It would be one in which social partner phenotypes are negatively correlated, so that an individual with the genes for the character of interest is differentially

likely to interact with a social partner that does *not* share those genes. Hamilton (1970) suggested that such a scenario would be conducive to the evolution of spite: by inflicting harm on their social partners, even at a cost to themselves, individuals with the genes for spite could increase the relative representation of these genes in the next generation. Hamilton's rule underwrites this prediction, since it implies that, when r is negative, a social behaviour can be favoured by selection even if its fitness effects on both actor and recipient are also negative (Gardner and West 2004a, b, 2006; West and Gardner 2010; Smead and Forber 2013).

These properties of the modern, statistical conception of relatedness are admittedly counter-intuitive. They reveal the extent to which relatedness-*qua*-technical-term has departed from our everyday, pre-theoretical notion of genealogical kinship. Yet we have good theoretical reasons to prefer these statistical measures of r to more traditional, pedigree-based measures, in spite of their counterintuitiveness. For when the pedigree-based and statistical measures part ways, it is the latter, not the former, that yield the right predictions about the response to selection when plugged into Hamilton's rule.

3 Gene Mobility as a Source of Genetic Similarity

Since the 1940s, microbiologists have known that prokaryotes are able to transfer genes 'horizontally' within a single generation, bypassing the parent-offspring channel (Lederberg and Tatum 1946; Zinder and Lederberg 1952). This horizontal gene transfer occurs most readily among conspecifics. Occasionally, genes cross species boundaries, and these events, though relatively rare, have profound implications for the notion of a 'tree of life' (O'Malley 2010; O'Malley and Boucher 2011). Here, I intend to focus mainly on within-species gene mobility, and on the implications it holds for our understanding of social evolution.

It is already widely acknowledged that gene transfer among conspecifics affects the course of evolution by providing material for homologous recombination (Frost et al. 2005). But more recent literature has highlighted another way in which gene transfer among conspecifics has consequences for social evolution in microbes (Nogueira et al. 2009; Rankin et al. 2011a, b; Mc Ginty et al. 2011; Mc Ginty and Rankin 2012; Mc Ginty et al. 2013; West and Gardner 2013). These consequences arise not from its role as a source of recombination, but from its role as a source of genetic similarity between nearby organisms.

I do not intend to discuss *all* forms of gene mobility in this article. 'Horizontal gene transfer' and 'gene mobility' are in fact umbrella terms for a plurality of different mechanisms, including transformation, transduction and conjugation (Bushman 2002; Thomas and Nielsen 2005). Similarly, 'mobile genetic element' is an umbrella term for various types of genetic entity that can be transferred in this way, including transposable elements, bacteriophages and plasmids (Frost et al. 2005). My

focus here will be on the evolutionary consequences of one particular mechanism, namely plasmid transfer by bacterial conjugation.

The process of conjugation is sometimes likened to microbial sex, presumably owing to the recombination in which it issues. It is not, however, a mode of reproduction: no cell division takes place. Instead, nearby cells come into contact via tubular protusions known as pili. Plasmids—packets of extra-chromosomal DNA—may move between adjacent cells via these pili. The mechanistic details need not concern us here. The crucial point about plasmid transfer by conjugation, for our purposes, is that it is a *replicative* process: the plasmid does not move from one cell to another, but is copied. The result is that, after conjugating, two nearby organisms may share genetic elements that they did not share before. We can therefore see how, in principle at least, plasmid conjugation can be a source of genetic similarity between social partners.

For a simple example, imagine a population sorted randomly into large groups of size M, in which some individuals carry a rare plasmid X. Because the process of group formation is random, there is initially no tendency for X-bearers to cluster together. Suppose, however, that at a particular point in the life-cycle, each individual conjugates with a randomly picked member of its group. Further suppose that, if a bearer conjugates with a non-bearer during this conjugation phase, the plasmid will infect the non-bearer with probability λ. After transfer, groups that had no plasmid bearers prior to the transfer phase will still have no plasmid bearers, but groups that had N plasmid bearers before the transfer phase are now expected to have approximately $N + \lambda N$ bearers.[3] The distribution of plasmid bearers across groups is now non-random: plasmid bearers now cluster together in a way they did not before. Models of this general sort will be discussed in greater detail below.

First, however, I must confront a sceptical worry: is conjugation really a *significant* source of genetic similarity from the point of view of inclusive fitness theory? In other words, is it ever likely to make a significant difference to the value of r? This will depend in part on the measure of r we employ. Pedigree-based measures of r will not be affected by mobile genetic elements, but generalized statistical measures of r may well be. The extent to which statistical measures of r are affected will depend on the extent to which mobile genetic elements are implicated in producing social phenotypes. For (as we saw above) these statistical measures do not take *every* genomic locus into account, but only those loci which are relevant to the social character under investigation. The upshot is that inclusive fitness theorists are entitled to ignore genetic correlations generated by conjugation unless they occur at loci that are relevant to the production of social phenotypes.

Nevertheless, recent empirical work suggests that inclusive fitness theorists often cannot afford

[3]This approximation relies on our assumptions that the plasmid is rare and that groups are large. If we do not make these assumptions, the frequency of plasmid bearers after conjugation is $N + \lambda N(M - N)/(M - 1)$. This quantity is approximated by $N + \lambda N$ when $M >> N$ (the plasmid is rare) and $M >> 1$ (groups are large). I thank Patrick Forber and an anonymous referee for helpful suggestions here.

to ignore mobile genetic elements, because these entities are indeed implicated in producing social behaviour in microbes—and more often than one might imagine. A study by Teresa Nogueira and colleagues (2009), which examined 21 *Escherichia coli* genomes, concluded that 'genes coding for secreted proteins—the secretome—are very frequently lost and gained and are associated with mobile elements' (2009, p. 1683). In a recent review, Daniel Rankin and colleagues (2011a) synthesize diverse sources of evidence in support of the hypothesis that mobile genetic elements are 'drivers of bacterial sociality' (2011a, p. 5).

What might explain this surprisingly strong association between social phenotypes and mobile genetic elements? As Rankin et al. observe, one possible explanation is that social phenotypes evolve more readily when they are encoded by mobile genes, owing to the effects of gene mobility on relatedness. The thought, in broad terms, is that the very mobility of these genes leads to positive genetic correlation between nearby organisms at relevant loci—and thereby leads to a population structure that is conducive to the evolution of cooperation. In a manner of speaking, we might say that mobile genetic elements, by generating genetic similarity between adjacent cells, are able to construct a social niche in which the cooperative phenotypes they encode are selectively advantageous.

4 A Temporal Aspect to Relatedness

The role of gene mobility in microbial social evolution is ultimately an empirical question; and, although the available evidence is suggestive, it remains an open one. I do not intend to settle it here. Instead, I want to take a step back from this empirical question to focus on a conceptual question. Suppose that gene mobility is indeed an important source of genetic similarity in microbial populations. *What follows for the concept of relatedness?*

One thing is immediately obvious: if gene mobility is a significant source of genetic similarity in microbial contexts, then pedigree-based measures of r will be inadequate in these contexts. This is because pedigree-based measures only take account of genetic similarity due to shared genealogy: they are insensitive to patterns of similarity that arise from other sources, even though these patterns may significantly alter the value of a recipient to an actor as a route to genetic representation in future generations. We should not, however, overstate the significance of this development. For as I emphasized in Section 2, inclusive fitness theorists already regard pedigree-based measures of r as deeply problematic, partly in virtue of the inability of these measures to accommodate greenbeard effects, but mostly in virtue of their tendency to mislead whenever selection is strong. Indeed, in light of these problems, many theorists have already switched to employing generalized statistical measures of r in formal work, and experimental biologists have already developed sophisticated empirical techniques for evaluating these measures in the field. These generalized measures of r make no assumptions about the mechanisms responsible for genetic similarity between organisms,

and so can, in principle, accommodate correlations generated by mobile genetic elements.

One might infer from this that inclusive fitness theory is 'preadapted' for gene mobility—that, for quite different reasons, theorists have already undertaken the conceptual revisions necessary to apply inclusive fitness theory in a microbial setting. This inference, however, would be a little too hasty. There is no doubt that the generalized statistical measures of r canvassed in Section 2 are a major improvement on traditional pedigree-based measures, and that they are better able to accommodate the effects of mechanisms (such as bacterial conjugation) that generate genetic similarity independently of shared genealogy. But I contend that even generalized statistical measures of r run into difficulties when the genes controlling social phenotypes are mobile. And I further contend that these difficulties force yet another revision to the concept of relatedness—a revision that takes the concept even further away from its intuitive roots as a measure of genealogical kinship.

The source of all the trouble is a simple observation about gene mobility. As we saw in Section 2, relatedness is most commonly defined in contemporary inclusive fitness theory as a regression coefficient describing the statistical association between actor and recipient genotypes, where genotypes are characterized quantitatively by genetic values. This definition takes it for granted that each organism in the population of interest can be unambigously assigned to a genotype (i.e. given a determinate genetic value) with respect to the trait of interest. But note that, when organisms are horizontally exchanging genes for social phenotypes at a non-negligible rate, we can no longer even talk of an organism's genotype or genetic value *simpliciter*, for these properties will tend to vary diachronically during the course of its life-cycle. Strictly speaking, we can only talk of an organism's genotype or genetic value *at a particular time*.

This observation forces a revision to the regression definition of relatedness. The regression definition (as stated in equation (2)) treats an organism's genotype as a time-independent property. But when mobile genetic elements are relevant to the evaluation of genotypes, genotypes become time-dependent: an organism's genetic value must be evaluated at a particular time in its life-cycle, and we need to take account of the fact that it may make a difference whether the evaluation takes place before or after a plasmid transfer event.[4]

5 The Need for Diachronic Measures of r

How, then, should we modify our definition of r to incorporate this time-dependence? One might initially think that a quick fix is available. For one might think all we need to do is re-write our

[4] Even without gene mobility, there is a sense in which the average genotype of a social *group* may be time-dependent, if the composition of this social group changes during the life-cycle; and this can introduce a form of time-dependence to measures of whole-group relatedness. Time-dependence of this sort (and its consequences for r) is discussed in Úbeda and Gardner 2012. My claim here is that diachronic variation in an individual's genotype during its life-cycle introduces a yet more radical sort of time-dependence, with further consequences for the concept of relatedness.

definition from equation (2) to explicitly indicate that, in a microbial setting, all genetic values are a function of time (t):

$$r(t) = \frac{\text{Cov}\left(g_i(t), \hat{g}_i(t)\right)}{\text{Var}\left(g_i(t)\right)} \tag{4}$$

This definition of r makes room for the fact that genotypes can vary during the life-cycle by explicitly stating that the relatedness in a population at any particular time is equal to the synchronic genetic similarity between actors and recipients at that time.

Note, however, that this move immediately presents us with a further question. Our modified definition of r implies that, when genotypes vary during the life-cycle, r also varies. As a consequence, the value of $r(t)B - C$ will depend on the precise time in the life-cycle at which r is evaluated. We are thus led to ask: at which stage (or stages) in the life-cycle do we need to evaluate $r(t)$, if we want to use $r(t)B - C$ to predict the overall direction of the response to selection?

One intuitive proposal is that what ultimately matters to the direction of selection is the value of $r(t)$ at the time at which the altruistic action is actually performed, and at which the payoffs B and C are actually conferred. Let us call this time in the life-cycle the *time of action* (t_A). In other words, then, the proposal is that $r(t_A)$ gives us the coefficient of relatedness we need in order to predict the response to selection. This suggestion sounds attractive on first hearing, but on closer inspection turns out to be problematic. For what happens if two organisms that are genetically uncorrelated at t_A *become* correlated later on in their life-cycles, by virtue of conjugating?

Consider, for example, a simple model of pairwise cooperation in which pairs of organisms (drawn at random from an infinite population) interact socially at time t_A, then subsequently conjugate at time t_C (for more detail regarding the model, see Appendix A). Suppose there is a mobile genetic element, X, such that bearers of X reliably perform an altruistic behaviour at t_A, and thereby confer a benefit B on their partner at a cost C to themselves. $r(t_A)$ will be zero in this scenario, because social partners are drawn at random: there is no assortment mechanism to ensure that the benefits of the behaviour caused by X fall differentially on other bearers of X. Consequently, if we use $r(t_A)B-C$ to calculate inclusive fitness, we will conclude that expressing X co-varies negatively with inclusive fitness, and hence that selection will act to lower the frequency of X.

This conclusion, however, is potentially misleading. For note that the fitness benefit conferred by the behaviour will not be 'cashed in' by the recipient until the end of its life-cycle, when (if it survives) it will reproduce through cell division. Let us call the end of the life-cycle time t_T. Suppose that, as a consequence of conjugation after t_A but prior to t_T, a fraction of altruists succeed in transferring X to the recipient of their altruistic act. As a result of this conjugation process, the fitness benefits conferred by X at t_A fall differentially (to some extent) on organisms who are *subsequently* carriers of X at the end of their life-cycle. So although there is no assortment between altruists at t_A, conjugation leads to a form of *diachronic* assortment between altruists: the benefits conferred by X

fall differentially on *future* bearers of that gene. If this diachronic assortment is strong enough, then encoding the altruistic behaviour may be selectively advantageous for the plasmid. That is to say, X may increase in frequency in virtue of the altruistic behaviour it causes, at the expense of otherwise identical mobile genes that have no effect on the behaviour of their bearers.

The implication is that $r(t_A)$ is not the optimal measure of relatedness to use in the model, if we want to predict the response to selection using Hamilton's rule. In general, if a social actor will have an opportunity to conjugate with a recipient after conferring a benefit upon it, then that recipient has value to the actor as a means of transmitting any plasmids it carries to future generations. Because $r(t_A)$ considers only synchronic genetic similarity at the time of action, it will take no account of this value, and may consequently mislead regarding the direction of selection.

So what *is* the best measure of relatedness in this model? One might suppose that the best measure would be $r(t_T)$, the synchronic genetic similarity among group members at the end of their life-cycle. After all, by t_T we can be sure that all conjugation has taken place, so we can be sure that the genotypes the organisms have at this time are their final genotypes—the genotypes they will transmit to the next generation, if they reproduce. Yet this measure is also problematic, because social groups that existed at t_A may have dispersed by this time, so any positive assortment that existed at t_A may have disappeared by t_T. For similar reasons, we cannot simply identify relatedness with the average synchronic correlation among group members over the course of the life-cycle. For in a scenario in which group members interact socially, momentarily conjugate and then immediately disperse, synchronic correlation may only exist very briefly, and so would make little difference to such an average; yet the genes for social phenotypes may still be able to spread.

A formal treatment of the model, given in Appendix A, shows that, if we want to use Hamilton's rule to predict whether or not it is evolutionarily advantageous for X to encode an altruistic phenotype, then the appropriate measure of r is in fact the following covariance ratio:

$$r = \frac{\mathrm{Cov}\left(g_i(t_T), \hat{g}_i(t_A)\right)}{\mathrm{Cov}\left(g_i(t_T), g_i(t_A)\right)} \tag{5}$$

This quantity is not a measure of synchronic genetic similarity between social partners at *any* single stage in the life-cycle, but rather a measure of *diachronic* genetic similarity between actors at t_A and recipients at t_T. In other words, it quantifies the degree of diachronic genetic similarity between *actors at the time of action* and *recipients at the end of their life-cycle*. This measure of r is influenced not only by pre-existing assortment between of action at the time of action, but also by diachronic assortment generated by any subsequent plasmid transfer. It thus provides a more accurate conception of the value of a recipient to an actor as a means of transmitting X to future generations.

The general conceptual point here is that, whenever actors will have the opportunity of transferring genes to their recipients after interacting socially with them, the recipient has value to the actor

as a route to genetic representation in future generations, and we should take this indirect transmission pathway into account when calculating the actor's inclusive fitness. In general, the best way to do this is to define r diachronically, as a measure of the genetic similarity between actors at the time of action and recipients at the end of their life-cycle.

This point is not restricted to our simple model: the same considerations apply in any model in which conjugation occurs after social interaction. The precise measure of r, however, will vary depending on the precise details of the model. Public goods games in group-structured populations provide a much more realistic model of microbial cooperation than games of pairwise interaction in an unstructured population. In Appendix B, I show that, in a public goods game in which conjugation occurs after public goods production, the best measure of r is the following covariance ratio:

$$r = \frac{\mathrm{Cov}\left(g_i(t_T), G_i(t_A)\right)}{\mathrm{Cov}\left(g_i(t_T), g_i(t_A)\right)} \tag{6}$$

Where G_i is the average genetic value of the focal individual's social group. This quantity measures the degree of diachronic genetic similarity between the composition of the i^{th} individual's social group at t_A and its own personal genotype at t_T. Again, therefore, it may be glossed informally as a measure of association between actors at the time of action and recipients at the end of their life-cycle. The difference is that $\hat{g}_i$, the genetic value of the focal individual's social partner, is replaced by G_i, the average genetic value of its social group. The resulting measure of relatedness can be regarded as a version of the whole-group relatedness (equation (3)) extended to accomodate gene mobility.

6 The Relationship between Synchronic and Diachronic Measures

There are two special cases in which our diachronic measures (5) and (6) collapse into synchronic measures. Both are cases in which, for one reason or another, an organism's genotype at the time of action (g_A) can be identified with its genotype at the end of its life-cycle (g_T) so that the diachronic correlation between social partners across this time interval is simply equal to the synchronic correlation between social partners evaluated at either of these times, or (for that matter) at any time in between.

The first (and most obvious) case is one in which there is no plasmid transfer at all. If there is no plasmid transfer, then organisms are stuck with their initial genotypes, and we can treat an organism's genotype as if it were time-independent. We therefore recover our traditional measures of 'others-only' and 'whole-group' relatedness (2) and (3) as special cases of the diachronic measures (5) and (6), respectively. This shows that it is entirely reasonable to use the standard measures of relatedness if the genes for the social behaviours we are interested in are not carried by mobile genetic

elements. It seems reasonable to assume that this is typically the case in populations of multicellular animals, though we should perhaps not take such things entirely for granted (cf. Dunning Hotopp et al. 2007).

The second case is one in which there is plasmid transfer, but all transfer of genetic material between individuals occurs prior to social interaction. In this case too we can identify g_A with g_T, and can hence identify r with the synchronic relatedness at either point in time, or at any point in between. This special case is investigated in detail by Mc Ginty et al. (2013). Their analysis is illuminating, and I have drawn on it in my own arguments (see Appendices), but in my view they misstep in suggesting that it is *only* in this special case that gene mobility can alter the value of r:

> Our results clearly depend on our life cycle and, if transmission were to occur after the public goods interaction [...] then we would no longer see the kin selection effect but the infectivity effect would remain. (2013, p. 6)

If this were true, it would be rather disappointing. After all, the assumption that *all* horizontal transfer occurs *before* any public goods production is unlikely to obtain in any real microbial population: as Mc Ginty et al. readily admit, it is an idealization made for the sake of simplicity. So if the impact of gene mobility on inclusive fitness could be wholly explained as an artefact of this idealizing assumption—if it could be shown to disappear as soon as the assumption is relaxed—then we would have no reason to expect to find any such effect in the real world, and no reason to think it could explain any real-world instances of microbial sociality. The scope of Mc Ginty and colleagues' conclusions would thus be severely limited.

Luckily, however, Mc Ginty et al. overstate the importance of their idealizing assumption. My analysis shows that there is still a 'kin selection' effect on the direction of social evolution (in a suitably broad sense of the word 'kin') even if plasmid transfer occurs partially or wholly after social interaction. It is simply that, to capture this effect, one must work with an unorthodox, diachronic measure of relatedness, to allow for the fact that the benefits of altruism may fall differentially not on current bearers of the social plasmid, but on future bearers.

7 Open Empirical Questions

7.1 Does altruism sometimes evolve without correlated interaction?

The focus of the foregoing discussion has been on a conceptual question: if (as seems likely) gene mobility has a significant impact on the evolution of cooperation in microbes, what follows for the concept of relatedness? I have argued that the best measure of relatedness in microbial contexts will often be diachronic—it will be a measure of the genetic similarity between actors at the time

of action and recipients at the end of their life-cycle. In essence, this diachronic measure is needed to account for the evolutionary consequences of what we might call 'ship jumping', in which the plasmids responsible for a social phenotype subsequently 'jump ship' to the recipients of the fitness benefits they induced their original host to donate.

But the possibility of such 'ship jumping' not only pushes us to rethink the concept of relatedness; it also threatens an important piece of received wisdom regarding the evolution of altruism. For it implies that it may, in principle, be evolutionarily advantageous for a plasmid to encode an altruistic phenotype (such as the production of a costly public good) even when there is no correlated interaction between altruists. The models analysed in the Appendix demonstrate this, since it is assumed in these models that groups are formed completely at random, with no positive assortment at the time of interaction—and yet it is shown that an altruistic phenotype can still increase in frequency, provided the plasmid that encodes it is subsequently able to generate diachronic assortment between actors and recipients through horizontal transfer.

Although this is an 'in principle' result—a proof of possibility, rather than an empirically substantiated claim—it is one with considerable potential significance, for it casts doubt on the commonly-held view (based, of course, on studies of multicellular animals) that correlated interaction between altruists is necessary for the evolution of altruism.[5] In microbes, the assortment that enables the evolution of altruism need not be there at the time of interaction; indeed, synchronic assortment at any time is unnecessary. The requisite assortment can be purely diachronic, and it can be generated by the plasmids themselves through transfer events that occur later on in their hosts' life-cycle, some time after social interaction has taken place.

Note that the point here is not simply that, in cases of uncorrelated interaction, a plasmid may spread *in spite* of the altruistic phenotype it produces in its host, by virtue of its underlying infectivity (cf. Smith et al. 2001; Giraud and Shykoff 2011). The point is rather that, in cases of uncorrelated interaction, 'ship jumping' can make encoding altruistic phenotypes *selectively advantageous* for a plasmid, so that a plasmid which encodes an altruistic phenotype spreads faster than it would have done had it been phenotypically inert.[6] The models analysed in the Appendix demonstrate the possibility of this phenomenon by explicitly comparing the change in frequency of an altruism-encoding plasmid to the change in frequency of one that is equally infective but neutral with regard to the fitness of its host. The models further demonstrate that the altruism-encoding plasmid will outperform the inert

[5] See, e.g., McElreath and Boyd 2007, 76: "The key to understanding the evolution of altruism is non-random interaction. If altruistic strategies are more likely to be paired with other altruistic strategies, altruism can evolve. If interaction is random, it cannot."

[6] One may ask: is the phenotype *really* altruistic, if it is advantageous for the plasmid that produces it? The answer is that it is altruistic in the standard technical sense of the term, provided the phenotype detracts from the fitness of the organism that expresses it and confers benefits on other organisms. In general, any altruistic phenotype may be recast as 'selfish' from the point of view of the genes responsible, but this does not preclude its being altruistic in the standard technical sense of the term (West et al. 2007b; Okasha 2013)

plasmid only if rB outweighs C, where r denotes the diachronic measure of relatedness introduced above.

No amount of purely theoretical work, however, could settle the question of the importance of this 'ship jumping' mechanism in nature. In particular: does plasmid transfer by conjugation actually occur frequently enough in real bacterial populations for the mechanism to work? In other words, can the value of r at the plasmid locus really get high enough for rB to outweigh C? I have not begun to address this question here. Clearly, it could only be answered by empirical studies of the actual values of r, B and C in real populations. What we have, at this stage, is a credible hypothesis in need of further empirical support.

7.2 Why are altruism-encoding plasmids tolerated by the rest of the genome?

A further question concerns the extent to which plasmid transfer leads to intragenomic conflict (Mc Ginty and Rankin 2012). After all, there is a sense in which our 'ship jumping' mechanism is a greenbeard mechanism: just as in a classic greenbeard scenario, a gene for altruism is able to spread by virtue of generating genetic correlation between social partners *at its own locus*, without thereby generating correlation at other genomic loci (e.g. chromosomal loci). The benefits of the altruism caused by the plasmid fall differentially on future bearers *of that plasmid*, potentially allowing it to increase in frequency, but there is no reason to expect that they fall differentially on recipients who are genetically similar to the actor at other loci.

We thereby intuitively expect that genes at other loci will evolve to inhibit the uptake and expression of the mobile genetic element. How easy this is in practice will, of course, depend on the mechanistic details. As Mc Ginty and Rankin (2012) point out, it may well be that an organism cannot easily block the uptake and expression of a plasmid that encodes an altruistic trait without also inhibiting the uptake and expression of other plasmids that are beneficial to their bearer. It may consequently be in the interests of the genome as a whole to tolerate a proportion of plasmids that encode altruistic traits in order to receive the benefits from other plasmids that are not so costly. Another possibility is that genes that dispose their bearers to produce public goods are sometimes linked on the same plasmid to genes that benefit their host, so that the net fitness effect of acquiring the plasmid is positive. However, this too remains an open empirical question.

8 Conclusion

This year, 2014, marks the 50th anniversary of Hamilton's original formulation of inclusive fitness theory. In the intervening decades, the theory has been repeatedly questioned and challenged. Successive generations of critics have sought to expose the limitations of the theory, while successive generations of inclusive fitness theorists have sought to extend Hamilton's original models to over-

come these alleged limitations (Bourke 2011a, b; Gardner et al. 2011; Birch forthcoming). What these protracted debates have shown is that Hamilton's central insight is far more general than even he initially supposed. What Hamilton saw was that altruism can evolve by natural selection whenever recipients provide actors with an indirect means of transmitting the genes for altruism to the next generation. Shared ancestry is one important mechanism by which such a situation can arise, but it is not the only such mechanism. We now know that greenbeard-type mechanisms, once thought to be the preserve of thought experiments, exist in nature. And, in microbial populations, a growing body of evidence suggests that bacterial conjugation (which can be viewed, from a certain oblique angle, as a species of greenbeard mechanism) cannot be discounted either.

I have argued that, when such effects are at work, relatedness must be rethought. This is because what matters for the evolution of social behaviour in these contexts is diachronic similarity between actors at the time of action and recipients at the end of their life-cycle. This diachronic assortment may in principle arise without any synchronic assortment (either at the time of action or at the end of the life-cycle) if the genes for altruism are able to 'jump ship' from actor to recipient after causing the altruistic act. There is much we do not yet understand about the relationship between microbial sociality and gene mobility, and we do not yet know how important such mechanisms will actually turn out to be. We can, however, draw the following conditional conclusion: if gene mobility does indeed turn out to be a significant influence on the evolution of microbial cooperation, then we will need to change the way we conceptualize relatedness.

Acknowledgements

Versions of this article were presented between November 2012 and January 2014 at the University of Cambridge, the London School of Economics and Political Science, Carnegie Mellon University, BSPS 2013 (at the University of Exeter) and ISHPSSB 2013 (at the University of Montpellier). I thank the audiences at these events for their questions and comments, and I particularly thank my co-panellists at ISHPSSB 2013, Sara Mitri and John Dupré. I thank Christopher Clarke, John Dupré, Patrick Forber, Andy Gardner, Tim Lewens, Conor Mayo-Wilson, Sara Mitri, Tobias Uller and an anonymous referee for their comments on the manuscript. I also extend belated thanks to Rosa Cao for her comments on another paper.

Appendix A: Diachronic Relatedness in a Prisoner's Dilemma

In the main text I argue that, when organisms engage in plasmid conjugation following pairwise social interaction, the best measure of r is a diachronic variant of the usual regression definition. This appendix supports that argument by providing a formal analysis of a simple model of pairwise

cooperation. This is not intended to provide a realistic model of microbial cooperation; it is a simple model designed to make the conceptual issues surrounding the definition of r as clear as possible.

Basic setup

We consider an infinite population of asexually reproducing organisms. The population is sorted at random into pairs, and each pair plays a one-shot two-player Prisoner's Dilemma with the following payoff matrix:

	COOPERATE	DEFECT
COOPERATE	$B - C$	$-C$
DEFECT	B	0

We assume that the population is under weak viability selection only, so that the payoffs can be interpreted as marginal costs and benefits to the viability of the players.

A fraction f_X of organisms bear a plasmid, X, that reliably causes them to cooperate. Let $g_{Xi}(t)$ represent the genic value of the ith organism with respect to X at time t, such that $g_{Xi}(t) = 1$ for an organism that carries X at t, and $g_{Xi}(t) = 0$ for an organism that does not. Meanwhile, a fraction f_Y of organisms bear a different plasmid, Y, that has no fitness-relevant phenotypic effects on its host. Let $g_{Yi}(t)$ represent the genic value of the i^{th} organism with respect to Y at time t, such that $g_{Yi}(t) = 1$ for an organism that carries Y at t, and $g_{Yi}(t) = 0$ for an organism that does not. Both X and Y may in principle be carried by the same organism. We assume, however, that the two plasmids are transmitted independently of each other, so that there is no correlation, positive or negative, between possessing X and possessing Y.

After playing the game and conferring payoffs, the players conjugate with probability π. During conjugation, a bearer of plasmid X will transmit the plasmid to its social partner (if it is initially a non-bearer) with transmission probability λ_X. The λ_X parameter may be interpreted as quantifying the infectivity of X. Correspondingly, a bearer of plasmid Y will transmit the plasmid to its social partner (if it is initially a non-bearer) with transmission probability λ_Y. The λ_Y parameter may be interpreted as quantifying the infectivity of Y.

Next, there is a global competition phase, in which all individuals compete for representation in the next generation, and individuals are killed off with a probability equal to their net viability (i.e. individuals who received B without paying C are least likely to be killed off; individuals who paid C without receiving B are most likely). Finally, at the end of the life-cycle, the surviving cells divide to produce the next generation. We assume that all plasmids are transmitted from a parent cell to its daughter cells with perfect fidelity. In reality, plasmids are sometimes 'lost' during cell division

(i.e. they are absent from one of the daughter cells). We neglect this phenomenon here, because it would complicate the analysis without affecting the main conceptual conclusions.

Because individuals are paired up at random in the social phase, there is no positive assortment at t_A. We thus intuitively expect that defectors will outperform cooperators, and that plasmids (such as X) that encode cooperation will decrease in frequency at the expense of those (such as Y) that do not. But, as we will see, the possibility of plasmid transfer after social interaction complicates the picture.

Modified Price equation

We can employ the Price equation (Price 1970, 1972) to study the conditions under which X will spread more rapidly than Y, or vice versa. The standard Price equation describes the change in the additive genetic value of some trait between an ancestor population and a descendant population. These populations may represent consecutive generations in a discrete generations model, or appropriately chosen earlier and later census points in an overlapping generations model. Price showed that this change could be decomposed into a covariance term and an expectation term:

$$\Delta\overline{g_i} = \frac{1}{\overline{w}}\left[\mathrm{Cov}(w_i, g_i) + \mathrm{E}(w_i \Delta g_i)\right] \tag{7}$$

Where w_i is the fitness of the i^{th} individual, conceptualized as its total contribution to the descendant population, and g_i is its additive genetic value for the character of interest.

The standard version of the Price equation is not ideal in the present context, for two main reasons. Firstly, the genetic value in the Cov term is not indexed to any particular time in the life-cycle, so its value will be indeterminate in cases in which an organism's genotype is altered during its life-cycle by plasmid transfer. For our purposes, we need to specify the time in the life-cycle at which g_i is to be evaluated. Secondly, both terms in the standard Price equation are affected by fitness differences; the equation thus splits the effects of differential fitness between its two terms. For our purposes, it will be convenient to capture all the effects of differential fitness in a single term. We overcome both these drawbacks by employing a modified version of the Price equation first derived by Frank (1998):

$$\Delta\overline{g_i} = \frac{1}{\overline{w}}\left[\mathrm{Cov}(w_i, g'_i)\right] + \mathrm{E}(\Delta g_i) \tag{8}$$

Where g'_i is the *transmitted* genetic value of the i^{th} individual with respect to the character of interest, i.e. the average genetic value of its descendants in the next generation.

In our model, in which we assume that plasmids are transmitted with perfect fidelity from parent cells to daughter cells, g'_i can be identified with $g_i(t_T)$, the terminal genetic value of the i^{th} individual

at the end of its life-cycle (that is, at either the time of cell death or the time of cell division, depending on the organism's fate). Δg_i can then be interpreted as the average change in an organism's genetic value between the beginning and end of its own life-cycle, while $\Delta\overline{g_i}$ can be interpreted as the change in the population mean of g during a single iteration of the life-cycle. Hence:

$$\Delta\overline{g_i} = \frac{1}{\overline{w}}\left[\mathrm{Cov}(w_i, g_i(t_T))\right] + \mathrm{E}(\Delta g_i) \tag{9}$$

This is a particularly useful version of the Price equation in a microbial context, and has been employed to this end in recent work by Mc Ginty et al. (2013).

Fitness function

In the Price formalism, the fitness of the i^{th} individual (w_i) is defined as its total contribution to the descendant population. In a discrete generations model, this is simply its total number of offspring. In our simple model, we can write this as a function of net viability, plus a residual component representing the deviation from expectation:

$$w_i = 2\left(V - C_i g_{Xi}(t_A) + B\hat{g}_{Xi}(t_A)\right) + \epsilon \tag{10}$$

As explained in the basic setup, the net viability of an organism depends on a baseline component plus the net payoff it receives in the social phase. This payoff in turn depends on the genetic values of itself and its social partner *at the time of action*; hence the genetic values in the fitness function should be evaluated at t_A. Note that every organism will deviate from expectation one way or the other, since it determinately either will or will not divide, and consequently its realized value for w_i will be either 2 or 0. This is unproblematic, provided the deviation from expectation does not co-vary with genetic value.

The intergenerational change in the frequency of X

We can derive an expression for the intergenerational change in the frequency of plasmid X by substituting equation (10) into equation (9), yielding:

$$\Delta\overline{g}_{Xi} = \frac{2}{\overline{w}}\left[-C\mathrm{Cov}(g_{Xi}(t_A), g_{Xi}(t_T)) + B\mathrm{Cov}(\hat{g}_{Xi}(t_A), g_{Xi}(t_T))\right] + \mathrm{E}(\Delta g_{Xi}) \tag{11}$$

Note that both V and ϵ have now dropped out of the analysis, because neither co-varies with genetic value. We can rearrange 11 as follows:

$$\Delta\overline{g}_{Xi} = \frac{2}{\overline{w}}\left[(rB - C)\,\mathrm{Cov}(g_{Xi}(t_A), g_{Xi}(t_T))\right] + \mathrm{E}(\Delta g_{Xi}) \tag{12}$$

Where r is identified with the following covariance ratio:

$$r = \frac{\mathrm{Cov}\left(g_{Xi}(t_T), \hat{g}_{Xi}(t_A)\right)}{\mathrm{Cov}\left(g_{Xi}(t_T), g_{Xi}(t_A)\right)} \tag{13}$$

We can therefore see that the expression for the change in frequency of X contains a Hamilton's rule-like component, albeit one that requires an unorthodox definition or r (more on this below). However, we can see that it also contains a separate, second component, $\mathrm{E}(\Delta g_{Xi})$, which reflects the fact that X may also spread partly in virtue of its infectivity, irrespective of its effects on its host's fitness (cf. Smith 2001; Giraud and Shykoff 2011).

The intergenerational change in the frequency of Y

The equation describing the change in frequency of plasmid Y is much simpler. Y has no fitness-relevant phenotypic effect on its host organism, and it is uncorrelated with the fitness-relevant plasmid X. Consequently, the presence or absence of Y does not co-vary with w_i. Hence $\mathrm{Cov}(w_i, g_{Yi}(t_T)) = 0$, and only the second term in our modified Price equation remains:

$$\Delta\overline{g}_{Yi} = \mathrm{E}(\Delta g_{Yi}) \tag{14}$$

This reflects the fact that Y is selectively neutral. If Y spreads, it can *only* be as a result of its infectivity.

When will X outperform Y?

By combining (12) with (14), we can obtain a condition under which the intergenerational change in the frequency of X is greater than the intergenerational change in the frequency of Y:

$$\Delta\overline{g}_{Xi} > \Delta\overline{g}_{Yi} \iff \frac{2}{\overline{w}}\left[(rB - C)\,\mathrm{Cov}(g_{Xi}(t_A), g_{Xi}(t_T))\right] > \mathrm{E}(\Delta g_{Yi}) - \mathrm{E}(\Delta g_{Xi}) \tag{15}$$

If we assume that X and Y are initially equal in their frequency ($f_X = f_Y$) and in their infectivity ($\lambda_X = \lambda_Y$), it follows that $\mathrm{E}(\Delta g_{Yi}) = \mathrm{E}(\Delta g_{Xi})$. This leaves the following simplified condition:

$$\Delta\overline{g}_{Xi} > \Delta\overline{g}_{Yi} \iff \frac{2}{\overline{w}}\left[(rB - C)\,\mathrm{Cov}(g_{Xi}(t_A), g_{Xi}(t_T))\right] > 0 \tag{16}$$

On the further but relatively mild assumption that $\mathrm{Cov}(g_{Xi}(t_A), g_{Xi}(t_T)) > 0$ (i.e. an organism's genotype at the time of action co-varies positively with its genotype at the end of its life-cycle), we can derive a condition under which X will outperform Y that is closely analogous to Hamilton's rule in its traditional form:

$$\Delta\overline{g}_{Xi} > \Delta\overline{g}_{Yi} \iff rB - C > 0 \tag{17}$$

This expression can be interpreted as a statement of the conditions under which a plasmid that encodes a social phenotype will outperform a plasmid that is identical in its infectivity but neutral with regard to its host's fitness. Importantly, the significance of this result is not limited to cases in which two very similar plasmids are actually competing. More generally, the expression can be interpreted as a statement of the conditions under which encoding a social phenotype is evolutionarily advantageous for a plasmid in a two-player Prisoner's Dilemma. If $rB - C < 0$, then the actual spread of the plasmid will be slower than the rate that would have been observed if the plasmid had been selectively neutral, so the plasmid gains no advantage by encoding the social phenotype. By contrast, if $rB - C > 0$, then the plasmid will spread more rapidly than an otherwise identical but selectively neutral plasmid would have done. In this situation, encoding the social phenotype is advantageous for the plasmid, despite the deleterious effect it has on the host.

The derivation of this condition could not proceed without the unorthodox definition of relatedness given in (13), on which r is defined a measure of the diachronic genetic similarity between actors at the time of action and recipients at the end of their life-cycle. Adopting this diachronic measure is therefore essential if we want to arrive at a Hamilton's rule-like condition that describes when encoding a social phenotype is evolutionarily advantageous for a plasmid in this model. Essentially, this is because in this model the advantageousness (or otherwise) of encoding the social phenotype is largely dependent on the probability ($\pi\lambda_X$) that the plasmid responsible will be able to 'jump ship' to the recipient after causing the altruistic act. Our diachronic measure of r, in contrast to more traditional measures, takes this 'ship jumping' phenomenon into account.

The dependence of r on $\pi\lambda_X$

To calculate the precise dependence of r on $\pi\lambda_X$, we begin by re-writing our covariance ratio as a ratio of regression coefficients:

$$r = \frac{\mathrm{Cov}\left(g_{Xi}(t_T), \hat{g}_i(t_A)\right)}{\mathrm{Cov}\left(g_{Xi}(t_T), g_{Xi}(t_A)\right)} = \frac{\beta_{\hat{g}_{Xi}(t_A), g_{Xi}(t_T)}}{\beta_{g_{Xi}(t_A), g_{Xi}(t_T)}} \tag{18}$$

We then compute the regression coefficients from the conditional expected values:

$$\beta_{\hat{g}_{Xi}(t_A), g_{Xi}(t_T)} = \mathrm{E}(\hat{g}_{Xi}(t_A)|g_{Xi}(t_T) = 1) - \mathrm{E}(\hat{g}_{Xi}(t_A)|g_{Xi}(t_T) = 0) = \pi\lambda_X \tag{19}$$

$$\beta_{g_{Xi}(t_A), g_{Xi}(t_T)} = \mathrm{E}(g_{Xi}(t_A)|g_{Xi}(t_T) = 1) - \mathrm{E}(g_{Xi}(t_A)|g_{Xi}(t_T) = 0) = 1 - (1 - f_X)\pi\lambda_X \tag{20}$$

This yields:

$$r = \frac{\pi\lambda_X}{1-(1-f_X)\pi\lambda_X} \tag{21}$$

From which we see that in this model, all else being equal, the diachronic assortment between altruists (r) increases with increasing probability of conjugation between social partners after social interaction (π) and with increasing probability of plasmid transmission when they conjugate (λ_X).

Appendix B: Diachronic Relatedness in a Public Goods Game

We turn now to a linear public goods game, a somewhat more realistic model of microbial cooperation. The analysis of this game largely parallels that of Appendix A. The aim is to derive a diachronic version of the whole-group relatedness (equation (3) in the main text) that provides the best measure of r in models in which plasmid transfer follows public goods production.

Basic setup

Consider, then, an infinite population of asexually reproducing organisms sorted randomly into groups of size N. As before, a fraction f_X of organisms bear a plasmid, X, that reliably causes them to produce a public good, while a fraction f_Y bear a different plasmid, Y, that has no fitness-relevant phenotypic effects on its host. Genetic values are defined as before.

At t_A, the organisms play a linear public goods game. Bearers of X produce a marginal viability benefit, B, that is shared evenly among the members of their social group. Each group member (including the producer) thus receives a total benefit $(G_{Xi}(t_A))B$, where $G_{Xi}(t_A)$ is the local (within-group) frequency of public good producers at t_A. Each producer incurs a viability cost C as a consequence of producing the public good.

As in the previous model, players may conjugate after the game. Each organism may conjugate at most once, and does so with probability π. These conjugating pairs are drawn at random from the group, and it is assumed that N is sufficiently large the frequencies of various possible pairs match the background frequencies in the global population. During conjugation, a bearer of plasmid X will transmit the plasmid to its social partner (if it is initially a non-bearer) with transmission probability λ_X. Correspondingly, a bearer of plasmid Y will transmit the plasmid to its social partner (if it is initially a non-bearer) with transmission probability λ_Y.

There follows a global competition phase, in which all individuals compete for representation in the next generation of groups, and individuals are killed off with a probability equal to their net viability. Finally, at the end of the life-cycle, the surviving cells divide to produce the next generation. We again assume that all plasmids are transmitted from a parent cell to its daughter cells with perfect fidelity.

Because groups are formed randomly, there is no positive assortment between the group members at t_A. We thus intuitively expect that public goods production will be undermined by free-riding, and that plasmids (such as X) that produce public goods will decrease in frequency at the expense of those (such as Y) that free-ride. Again, however, we find that allowing for the possibility of plasmid transfer after social interaction complicates the picture.

Equations for gene frequency change

As before, we take an appropriately rearranged Price equation as our starting point:

$$\Delta\overline{g_i} = \frac{1}{\overline{w}}\left[\mathrm{Cov}(w_i, g_i(t_T))\right] + \mathrm{E}(\Delta g_i) \tag{10}$$

The fitness function for this game is as follows:

$$w_i = 2\left(V - C_i g_{Xi}(t_A) + BG_{Xi}(t_A)\right) + \epsilon \tag{22}$$

Substituting (22) into (9), we obtain:

$$\Delta\overline{g}_{Xi} = \frac{2}{\overline{w}}\left[-C\mathrm{Cov}(g_{Xi}(t_A), g_i(t_T)) + B\mathrm{Cov}(G_{Xi}(t_A), g_{Xi}(t_T))\right] + \mathrm{E}(\Delta g_{Xi}) \tag{23}$$

Which can be re-written as:

$$\Delta\overline{g}_{Xi} = \frac{2}{\overline{w}}\left[(r_G B - C)\,\mathrm{Cov}(g_{Xi}(t_A), {}_{Xi}(t_T))\right] + \mathrm{E}(\Delta g_{Xi}) \tag{24}$$

Where r_G is now identified with the following covariance ratio:

$$r_G = \frac{\mathrm{Cov}\left(g_{Xi}(t_T), G_{Xi}(t_A)\right)}{\mathrm{Cov}\left(g_{Xi}(t_T), g_{Xi}(t_A)\right)} \tag{25}$$

Plainly, this covariance ratio is very similar to that given in equation (13); the only difference is that a whole-group genetic value has replaced the social partner's genetic value in the numerator. The relationship between these covariance ratios is thus closely analogous to the relationship between the 'whole-group relatedness' (3) and the standard 'others-only' regression measure of r (2). We can think of the above ratio as a diachronic variant of the whole-group relatedness, designed to accommodate the effects of gene mobility occuring after social interaction.

Plasmid Y, as before, is selectively irrelevant, so its change in frequency is again given simply by:

$$\Delta\overline{g}_{Yi} = \mathrm{E}(\Delta g_{Yi}) \tag{26}$$

When will X outperform Y?

From here on, the analysis is virtually identical to that of Appendix A. Combining (24) with (26), we see that X outperforms Y in this model under the following conditions:

$$\Delta\overline{g}_{Xi} > \Delta\overline{g}_{Yi} \iff \frac{2}{\overline{w}}\left[(r_G B - C)\,\mathrm{Cov}(g_{Xi}(t_A), g_{Xi}(t_T))\right] > \mathrm{E}(\Delta g_{Yi}) - \mathrm{E}(\Delta g_{Xi}) \quad (27)$$

In the special case in which plasmids X and Y are equal in their frequency and in their degree of infectivity, and on the further assumption that $\mathrm{Cov}\left(G_{Xi}(t_T), G_{Xi}(t_A)\right) > 0$ (i.e. a social group's average genetic value at t_A co-varies positively with its average genetic value at the end of the life-cycle), we can again derive a Hamilton's rule-like condition under which X will be selectively favoured over Y:

$$\Delta\overline{g}_{Xi} > \Delta\overline{g}_{Yi} \iff r_G B - C > 0 \quad (28)$$

This expression can be interpreted as a statement of the conditions under which encoding public goods production is evolutionarily advantageous for a plasmid in a public goods game. If $r_G B - C < 0$, then the actual spread of the plasmid will be slower than the rate that would have been observed if the plasmid had been selectively neutral, so the plasmid gains no advantage by encoding the social phenotype. By contrast, if $r_G B - C > 0$, then the plasmid will spread more rapidly than an otherwise identical but selectively neutral plasmid would have done. In this situation, producing the public good is advantageous for the plasmid, despite the deleterious effect it has on the host.

As before, the derivation of this result relies on our adopting an unorthodox, diachronic definition of r. This time, r quantifies the fraction of the overall diachronic covariance between genotypes at t_A and genotypes at t_T that can be accounted for by covariance between the earlier and later group genetic values. It is a natural extension of the 'whole-group relatedness' to the diachronic case, and this extension is needed in order to capture the diachronic genetic correlations between group members generated after the social phase by plasmid transfer.

The dependence of r_G on $\pi\lambda_X$

When (as in our model) all groups are of equal size N, there is a close relationship between the whole-group relatedness and the standard ('others-only') regression definition (Pepper 2000):

$$r_G = \frac{1 + (N-1)r}{N} \quad (29)$$

The same relationship holds between our modified diachronic measures for the whole-group and standard ('others-only') relatedness (i.e. (25) and (13) respectively).

In our model, each organism conjugates once (at most) with a randomly selected member of

its randomly formed group, and N is sufficiently large that the frequencies of the various types of pair in each group match the background frequencies in the population. Consequently, we can treat these pairs as if they had been drawn at random from the global population. The probability that they conjugate successfully is π, and the probability that an X-bearer will transfer X in this event is λ_X. Given these assumptions, the value of r will be the same as it was in the previous game. We can therefore combine (21) with (29) to obtain an expression for r_G:

$$r_G = \frac{1 + (N-1)(\pi\lambda_X/(1 - \pi\lambda_X(1 - f_X)))}{N} \tag{30}$$

This shows that, all else being equal, increasing π or λ_X increases the diachronic whole-group relatedness at the plasmid locus.

References

Berleman, J. E., T. Chumley, P. Cheung and J. R. Kirby. 2006. Rippling is a predatory behavior in *Myxococcus xanthus*. *Journal of Bacteriology* 188:5888-895.

Biernaskie, J. M., S. A. West and A. Gardner. 2011. Are greenbeards intragenomic outlaws? *Evolution* 65:2729-2742.

Birch, J. forthcoming. Hamilton's rule and its discontents. *British Journal for the Philosophy of Science*.

Bonner, J. T. 1959. *The cellular slime molds*. Princeton, NJ: Princeton University Press.

Bourke, A. F. G. 2011a. *Principles of social evolution*. Oxford: Oxford University Press.

Bourke, A. F. G. 2011b. The validity and value of inclusive fitness theory. *Proceedings of the Royal Society of London B: Biological Sciences 278:3313-3320*.

Breden, F. 1990. Partitioning of covariance as a method for studying kin selection. *Trends in Ecology and Evolution* 5:224-228.

Bushman, F. 2002. *Lateral DNA transfer: mechanisms and consequences*. Coldspring Spring Harbor, NY: Cold Spring Harbor Laboratory Press.

Cockerham, C. C. 1954. An extension of the concept of partitioning hereditary variance for analysis of covariances among relatives when epistasis is present. *Genetics* 39:8659-882.

Cornforth, D. M., D. J. Sumpter, S. P. Brown, and Å. Brännström. 2012. Synergy and group size in microbial cooperation. *American Naturalist* 180:296-305.

Crespi, B. J. 2001. The evolution of social behaviour in microorganisms. *Trends in Ecology and Evolution* 16:178-183.

Davies, N. B., J. R. Krebs and S. A. West. 2012. *An introduction to behavioural ecology*. Hoboken, NJ: Wiley-Blackwell.

Dawkins, R. 1976. *The selfish gene*. New York: W. W. Norton and Company.

Dunning Hotopp, J. C., M. E. Clark, D. C. Oliveira, J. M. Foster, P. Fischer, M. C. Muñoz Torres, J. D. Giebel, N. Kumar, N. Ishmael, S. Wang, J. Ingram, R. V. Nene, J. Shepard, J. Tomkins, S. Richards, D. J. Spiro, E. Ghedin, B. E. Slatko, H. Tettelin, J. H. Werren. 2007. Widespread lateral gene transfer from intracellular bacteria to multicellular eukaryotes. *Science* 317:1753-1756.

Dupré, J. 2012. *Processes of life: essays in the philosophy of biology*. New York: Oxford University Press.

Falconer, D. S. 1961. *Introduction to quantitative genetics (1st edition)*. Edinburgh: Oliver and Boyd.

Falconer, D. S. and T. F. C. Mackay. 1996. *Introduction to quantitative genetics (4th edition)*. London: Longman.

Fisher, R. A. 1918. The correlation between relatives on the supposition of Mendelian inheritance. *Transactions of the Royal Society of Edinburgh* 52:399-433.

Fisher, R. A. 1930. *The genetical theory of natural selection (1st edition)*. Oxford: Clarendon Press.

Fisher, R. A. 1941. Average excess and average effect of a gene substitution. *Annals of Human Genetics* 11:53-63.

Frank, S. A. 1997. Models of symbiosis. *American Naturalist* 150:S80-S99.

Frank, S. A. 1998. *Foundations of social evolution*. Princeton, NJ: Princeton University Press.

Frank, S. A. 2006. Social selection. In C. W. Fox and J. B. Wolf (eds.), *Evolutionary genetics: concepts and case studies*. Oxford University Press, pp. 350-363.

Frost, L. S., R. Leplae, A. O. Summers and A. Toussaint. 2005. Mobile genetic elements: the agents of open source evolution. *Nature Reviews Microbiology* 3:722-732.

Gardner, A. and K. R. Foster. 2008. The evolution and ecology of cooperation—history and concepts. In J. Korb and J. Heinze (eds), *Ecology of social evolution*. Heidelberg: Springer-Verlag, pp. 1-36.

Gardner, A. and S. A. West. 2004a. Spite and the scale of competition. *Journal of Evolutionary Biology* 17:1195-1203.

Gardner, A. and S. A. West. 2004b. Spite among siblings. *Science* 305:1413-1414.

Gardner, A. and S. A. West. 2006. Spite. *Current Biology* 16:R662-R664.

Gardner, A. and S. A. West. 2010. Greenbeards. *Evolution* 64:25-38.

Gardner, A., S. A. West and G. Wild. 2011. The genetical theory of kin selection. *Journal of Evolutionary Biology* 24:1020-1043.

Giraud, T. and J. A. Shykoff. 2011. Bacterial controlled by mobile elements: kin selection versus infectivity. *Heredity* 107:277-278.

Grafen, A. 1985. A geometrical view of relatedness. *Oxford Surveys in Evolutionary Biology* 2:28-89.

Grafen, A. 2006. Optimization of inclusive fitness. *Journal of Theoretical Biology* 238:541-63.

Haldane, J. B. S. 1955. Population genetics. In M. L. Johnson, M. Abercrombie and G. E. Fogg (eds), *New Biology 18*. London: Penguin, pp. 34-51.

Haldane, J. B. S. and S. D. Jayakar. 1962. An enumeration of some human relationships. *Journal of Genetics* 58:81-107.

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

Hamilton, W. D. 1964. The genetical evolution of social behaviour. *Journal of Theoretical Biology* 7:1-52.

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 J. F. Eisenberg and W. S. Dillon (eds), *Man and beast: comparative social behavior.* Washington, DC: Smithsonian Press, pp. 57-91.

Hamilton, W. D. 1972. Altruism and related phenomena, mainly in social insects. *Annual Review of Ecology and Systematics* 3:193-232.

Hamilton, W. D. 1975. Innate social aptitudes of man: an approach from evolutionary genetics. In R. Fox (ed.), *Biosocial anthropology*. New York: Wiley, pp. 133-55.

Hardin, R. 1982. *Collective action*. Baltimore, MD: John Hopkins University Press.

Kagel, J. H. and A. E. Roth (eds). 1995. *The handbook of experimental economics*. Princeton, NJ: Princeton University Press.

Kempthorne, O. 1954. The correlation between relatives in a random mating population. *Proceedings of the Royal Society of London B: Biological Sciences* 143:103-113.

Kempthorne, O. 1955. The theoretical values of correlations between relatives in random mating populations. *Genetics* 40:153-167.

Kempthorne, O. 1957. *An introduction to genetic statistics*. Oxford: Wiley.

Lederberg, N. D. and J. Tatum. 1946. Genetic exchange in Salmonella. *Journal of Bacteriology* 64:679-699.

Lewis, D. 1970. How to define theoretical terms. *Journal of Philosophy* 67:427-446.

Li, C. C. and L. Sacks. 1954. The derivation of joint distribution and correlation between relatives by the use of stochastic matrices. *Biometrics* 10:347-360.

McElreath, R. and R. Boyd. 2007. *Mathematical models of social evolution: a guide for the perplexed*. Chicago IL: University of Chicago Press.

Mc Ginty S. É., D. J. Rankin and S. P. Brown. 2011. Horizontal gene transfer and the evolution of bacterial cooperation. *Evolution* 65: 21-32.

Mc Ginty, S. É. and D. J. Rankin. 2012. The evolution of conflict resolution between plasmids and their bacterial hosts. *Evolution* 66:1662-1670.

Mc Ginty, S. É., L. Lehmann, S. P. Brown and D. J. Rankin. 2013. The interplay between relatedness and horizontal gene transfer drives the evolution of plasmid-carried public goods. *Proceedings of the Royal Society of London B: Biological Sciences* 280:20130400.

Marshall, J. A. R. 2011. Group selection and kin selection: formally equivalent approaches. *Trends in Ecology and Evolution* 26:325-332.

Maynard Smith, J. 1964. Group selection and kin selection. *Nature* 200:1145-1147.

Maynard Smith, J. 1976. Letter to New Scientist. *New Scientist* 71:247..

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

Nogueira T., D. J. Rankin, M. Touchon, F. Taddei, S. P. Brown and E. P. C. Rocha. 2009. Horizontal gene transfer of the secretome drives the evolution of bacterial cooperation and virulence. *Current Biology* 19(20):1683-1691.

Olson, M. 1965. *The logic of collective action: public goods and the theory of groups*. Cambridge, MA: Harvard University Press.

Okasha, S. 2013. Biological altruism. In E. Zalta (ed), *The Stanford encyclopedia of philosophy (Fall 2013 edition)*. URL=<http://plato.stanford.edu/archives/fall2013/entries/altruism-biological/> (Accessed 18/03/2014).

O'Malley, M. (ed.). 2010. The tree of life [special issue]. *Biology and Philosophy* 25(4).

O'Malley, M. (ed.). 2013. Philosophy and the microbe [special issue]. *Biology and Philosophy* 28(2).

O'Malley, M. and Y. Boucher (eds). 2011. Beyond the tree of life [online special issue]. *Biology Direct* 25(4).

Orlove, M. J. and C. L. Wood. 1978. Coefficients of relationship and coefficients of relatedness in kin selection: a covariance form for the RHO formula. *Journal of Theoretical Biology* 73: 679-686.

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

Price, George R. 1970. Selection and covariance. *Nature* 227:520-1.

Price, George R. 1972. Extension of covariance selection mathematics. *Annals of Human Genetics* 35:485-90.

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

Queller, D. C. 1985. Kinship, reciprocity, and synergism in the evolution of social behaviour. *Nature* 318:366-7.

Queller, D. C. and K. F. Goodnight. 1989. Estimating relatedness using genetic markers. *Evolution* 43: 258-275.

Rankin, D. J., E. P. C. Rocha and S. P. Brown. 2011a. What traits are carried on mobile genetic elements, and why? *Heredity* 106:1-10.

Rankin, D. J., S. E. Mc Ginty, T. Nogueira, M. Touchon, F. Taddei, E. P. C. Rocha and S. P. Brown. 2011b. Bacterial cooperation controlled by mobile genetic elements: kin selection and infectivity are part of the same process. *Heredity* 107:279-281.

Ridley, M. and A. Grafen. 1981. Are green beard genes outlaws? *Animal Behaviour* 29:954– 955.

Parsek, M. R. and E. P. Greenberg. 2005. Sociomicrobiology: the connections between quorum sensing and biofilms. *Trends in Microbiology*. 13:27-33.

Seger, J. 1981. Kinship and covariance. *Journal of Theoretical Biology* 91:191-213.

Smead, R. and P. Forber. 2013. The evolutionary dynamics of spite in finite populations. *Evolution* 67:698-707.

Smith, J. 2001. The social evolution of bacterial pathogenesis. *Proceedings of the Royal Society of London B: Biological Sciences* 268: 61-69.

Smith, J., J. D. van Dyken and P. C. Zee. 2010. A generalization of Hamilton's rule for the evolution of microbial cooperation. *Science* 328:1700-1703.

Strassmann, J. E. and D. C. Queller. 2011. Evolution of cooperation and control of cheating in a

social microbe. *Proceedings of the National Academy of Sciences USA* 108:10855-10862.

Strassmann, J. E., Y. Zhu and D. C. Queller. 2000. Altruism and social cheating in the social amoeba Dictyostelium discoideum. *Nature* 408:965-967.

Thomas, C. M. and K. M. Nielsen. 2005. Mechanisms of, and barriers to, horizontal gene transfer between bacteria. *Nature Reviews Microbiology* 3:711-721.

Toro, M., R. Abugov, B. Charlesworth and R. E. Michod. 1982. Exact versus heuristic models of kin selection. *Journal of Theoretical Biology* 97:699-713.

Úbeda, F. and A. Gardner. 2012. Genomic imprinting in the social brain: elders. *Evolution* 66:1567-81.

Uyenoyama, M. K. and M. W. Feldman. 1981. On relatedness and adaptive topography in kin selection. *Theoretical Population Biology* 19:87-123.

Velicer, G. J. and M. Vos. 2009. Sociobiology of the myxobacteria. *Annual Review of Microbiology* 63:599-623.

Wenseleers, T., A. Gardner and K. R. Foster. 2010. Social evolution theory: a review of methods and approaches. In T. Székely, A. J. Moore and J. Komdeur (eds), *Social behaviour: genes, ecology and evolution*. Cambridge: Cambridge University Press, 132-58.

West, S. A. and A. Gardner. 2010. Altruism, spite and greenbeards. *Science* 327:1341-1344.

West, S. A. and A. Gardner. 2013. Inclusive fitness and adaptation. *Current Biology* 23:R577-R584.

West, S. A., A. S. Griffin and A. Gardner. 2007b. Social semantics: altruism, cooperation, mutualism, strong reciprocity and group selection. *Journal of Evolutionary Biology* 20:415-432.

West, S. A., A. S. Griffin, A. Gardner and S. P. Diggle. 2006. Social evolution theory for microbes. *Nature Reviews Microbiology* 4:597-607.

West, S. A., S. P. Diggle, A. Buckling, A. Gardner and A. S. Griffin. 2007a. The social lives of microbes. *Annual Review of Ecology, Evolution and Systematics* 38:53-77.

Wright, S. 1922. Coefficients of inbreeding and relationship. *American Naturalist* 56: 330-338.

Zinder, N. D. and J. Lederberg. 1952. Genetic exchange in Salmonella. *Journal of Bacteriology* 64:679–699.