Adaptation and Inclusive Fitness

Review

Stuart A. West[1,*] and Andy Gardner[1,2]

Inclusive fitness theory captures how individuals can influence the transmission of their genes to future generations by influencing either their own reproductive success or that of related individuals. This framework is frequently used for studying the way in which natural selection leads to organisms being adapted to their environments. A number of recent papers have criticised this approach, suggesting that inclusive fitness is just one of many possible mathematical methods for modelling when traits will be favoured by natural selection, and that it leads to errors, such as overemphasising the role of common ancestry relative to other mechanisms that could lead to individuals being genetically related. Here, we argue that these suggested problems arise from a misunderstanding of two fundamental points: first, inclusive fitness is more than just a mathematical 'accounting method' — it is the answer to the question of what organisms should appear designed to maximise; second, there is something special about relatedness caused by common ancestry, in contrast with the other mechanisms that may lead to individuals being genetically related, because it unites the interests of genes across the genome, allowing complex, multigenic adaptations to evolve. The critiques of inclusive fitness theory have provided neither an equally valid answer to the question of what organisms should appear designed to maximise, nor an alternative process to unite the interest of genes. Consequently, inclusive fitness remains the most general theory for explaining adaptation.

Introduction

The most striking fact about living organisms is the extent to which they appear designed or adapted for the environments in which they live (Figure 1) [1]. The theory of natural selection provides an explanation. Darwin [2] pointed out that those heritable characters that are associated with greater reproductive success will tend to accumulate in biological populations, and he argued that this will lead organisms to appear as if they were designed to maximise their reproductive success. Hence, natural selection explains the appearance of design without invoking an intelligent designer.

More generally, the currently accepted paradigm for the study of adaptation, in fields such as animal behaviour, evolutionary ecology and sociobiology, is that organisms should appear designed to maximise their inclusive fitness, rather than their reproductive success [3–9]. Inclusive fitness captures how individuals are able to influence the transmission of their genes to future generations — they can either influence their own reproductive success (direct fitness) or the reproductive success of other individuals with which they share genes (indirect fitness) [10,11].

However, a number of papers [12–20] have questioned the generality and usefulness of inclusive fitness theory. In

particular, it has been argued that inclusive fitness is just one of many possible mathematical accounting methods for modelling the outcome of natural selection; that it is less correct than other mathematical methods for modelling when traits will be favoured by natural selection; and that it overemphasises the importance of common descent (kinship) relative to other mechanisms that can lead to individuals being genetically related [12–20]. If valid, these criticisms would have paradigm-shifting implications and textbooks would need to be rewritten.

Here, we show that controversy has arisen because four separate questions have been conflated (Table 1). Our aim is not to question the novelty or validity of the mathematical arguments that have been used to criticise inclusive fitness theory, which has already been done elsewhere [21–23]. Instead, our aim is to show that the critiques of inclusive fitness theory have asked the wrong questions, and hence are irrelevant, even before considering the mathematical arguments. Before we can address the different questions, we first need to explain the different approaches that have been used to conceptualise natural selection, and that are at the heart of this controversy.

Different Ways of Carving Up Natural Selection

The key criterion for natural selection to favour any trait is that the genes for this trait are positively correlated with individual fitness [24,25]. However, correlation need not imply causation, particularly when considering traits that can influence the fitness of other individuals which share genes for that trait, such as altruistic helping behaviours [10]. Consequently, social evolution researchers have found it helpful to partition natural selection into more meaningful, causal relations, when examining why traits are correlated with fitness. Debate has focused on the relative usefulness of three basic approaches: neighbour-modulated fitness, inclusive fitness and group selection (Supplemental information).

First, the 'neighbour modulated' or 'personal' fitness approach decomposes the overall correlation between an individual's genes and its fitness into a direct and an indirect effect (Figure 2A) [8,10,26–28]. The direct effect describes the impact of an individual's own genes on reproductive success. The indirect effect describes the impact of the genes carried by social partners of the focal individual upon the individual's own reproductive success. The theory of 'indirect genetic effects' takes a similar approach, but with a more explicit focus on phenotypic effects [29].

Second, the 'inclusive fitness' approach uses the same basic partition of natural selection into direct and indirect effects, but conceptualises the indirect effect in a different way (Figure 2B) [10]. Here, the indirect effect describes the impact of the focal individual's genes on the fitness of its social partners, weighted by genetic relatedness. Both the neighbour-modulated and inclusive fitness approaches lead to Hamilton's rule [10,11], which states that a trait will be favoured by natural selection when $-c + br > 0$ (where $-c$ is the direct fitness cost of the trait, b is the benefit provided to social partners by the trait and r is the genetic relatedness between the focal individual and its social partners).

The neighbour-modulated and inclusive fitness approaches differ in how they conceptualise the benefit (b)

[1]Department of Zoology, University of Oxford, The Tinbergen Building, South Parks Road, Oxford, OX1 3PS, UK. [2]Balliol College, University of Oxford, Broad Street, Oxford OX1 3BJ, UK.
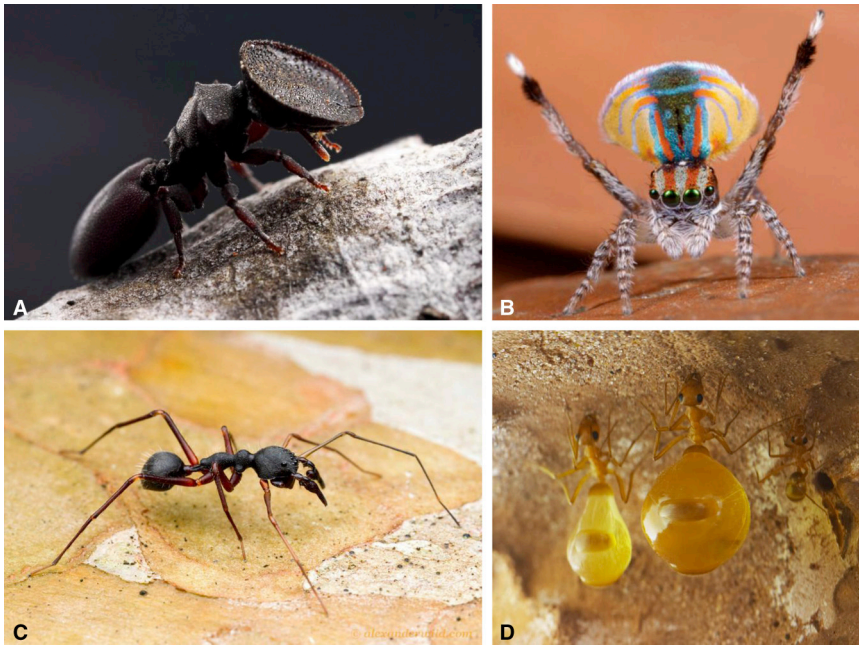*E-mail: stuart.west@zoo.ox.ac.uk

Figure 1. Organisms appear designed or adapted to the environment in which they live.

(A) A turtle ant soldier has a disc on its head that appears as if it was designed to fill the entrance to the tree cavities in which they live, forming a 'living door'. (B) A male peacock spider woos a female using a brightly-coloured flap that looks like an artist's psychedelic rendering of a spider. (C) This is not an ant — it is a clubionid spider that appears designed to look like the painfully-stinging trap jaw ants of the genus *Odontomachus*, which it lives alongside. (D) Some workers of the honeypot ant appear as if they are designed to act as storage containers. The cardinal problem for evolutionary theory is to explain this apparent design of organisms [5,57,64]. Photos used with permission, copyright Alex Wild (A,C,D) and Jürgen Otto (B).

and relatedness (*r*) terms [8,10,26–28]. For example, consider the evolution of altruism in a panmictic outbreeding population. With neighbour-modulated fitness, there is a tendency *r* that a focal individual who carries genes for altruism is also aided by her social partner. To the extent that she is aided, she receives a benefit *b*, and so the average benefit to the focal individual is *rb*. In contrast, with inclusive fitness, the focal individual always helps the social partner, providing a benefit *b*. In this case, *r* is a measure of how much helping the social partner increases the frequency of the focal individual's genes, such that we can express the overall genetic benefit to the focal individual as *rb*.

Consequently, neighbour-modulated fitness examines how social partners influence the fitness of the focal individual, whereas inclusive fitness examines how the focal individual influences the fitness of its social partners. The *b* term is the benefit that is either received (neighbour-modulated fitness) or given (inclusive fitness) by the focal individual. The *r* term is a measure of the extent to which either social partners have a similar disposition for altruism (neighbour-modulated fitness), or the focal individual values its social partners (inclusive fitness). Put simply, inclusive fitness is actor-centric with one focal actor, whom we focus on in a world full of recipients (including itself), while neighbour-modulated fitness is recipient-centric, with one focal recipient whom we focus on in a world full of actors (including itself).

Third, the 'group selection' approach decomposes the overall correlation between an individual's genes and its fitness into within-group and between-group effects (Figure 2C) [30]. The within-group effect describes the average association between genes and fitness within social groups in the population, and it contributes to the total action of natural selection in proportion to the heritable variation that exists within social groups. The between-group effect describes the association between genes and fitness at the group level, and contributes to the total action of natural selection in proportion to the heritable variation that exists between social groups.

Do Different Methods Yield Different Predictions?

Our first question is whether the different methods described above yield different predictions for how traits will evolve in response to natural selection? They do not [8,10,28,31–36]. They are just different ways of dividing up the dynamics of natural selection (Supplemental information). No approach is more correct than the other. Consequently, when modelling specific traits, researchers can use whatever method is most useful for the task at hand. The most frequently used methods are neighbour-modulated and inclusive fitness. In particular, modern neighbour-modulated fitness methods allow the modeller to go from the underlying biology to an expression or fitness, in a way that facilitates the development of relatively general models [8,27,28,35,37]. In contrast, the group selection approach is used relatively little for modelling specific traits partly because as soon as one moves away from the simplest, most abstract models, and wants to add in real world biology, it often becomes analytically intractable — for example, when populations are structured into different classes of individual, according to sex, age, caste or ploidy [38–40].

What Should Organisms Appear Designed for?

Our second question is which quantity does natural selection lead organisms to appear designed to maximise (design objective)? To put this another way: if natural selection leads organisms to appear as if they are striving to maximise their fitness, then what is their fitness [10,24,41]? Fitness, in the sense of the organism's maximand, must fulfil two criteria (Table 2).

First, the maximand must be a 'target' of natural selection. That is, the maximand must satisfy the condition that if a gene is positively correlated with this maximand, then natural selection will lead to that gene increasing in frequency [25,42]. This is true for neighbour-modulated fitness, and it is also true for inclusive fitness [10,24,41,43]. In contrast, it is not generally true that group fitness is a target of natural selection, because whether or not traits are favoured depends upon the relative importance of within-group and between-group fitness effects [40]. Consequently, not all genes that are associated with greater group fitness will

Table 1. Critiques of inclusive fitness theory have conflated separate questions.

Question	Answer
Do the different partitions of natural selection yield different predictions?	No.
What quantity does natural selection lead organisms to appear designed to maximise (design objective)?	Inclusive fitness.
Why is it useful to have a design principle or maximand?	To describe how organisms should be designed, and for linking theoretical and empirical studies.
Is there anything special about genetic relatedness that arises from recent common ancestry, as opposed to other genetic assortment mechanisms?	Yes, it leads to a more-or-less equal relatedness across most of the genome, and hence allows complex multigenic adaptations to be constructed.

be favoured by natural selection, and genes that are associated with lower group fitness can be favoured by natural selection.

Second, the maximand must be under the organism's 'full control', meaning that it is determined only by the traits and actions of the focal organism'. This is because organisms can only appear designed to maximise something that they are able to control. The individual does not, in general, have full control of its neighbour-modulated fitness, as parts of this are mediated by the actions of her social partners. However, the individual does have full control of inclusive fitness, as this is explicitly defined in terms of the fitness consequences for itself and others that arise out of its actions (Figure 2; Supplemental information) [10]. Specifically, Hamilton [10] described inclusive fitness as personal (neighbour-modulated) fitness, stripped of all components caused by others (leaving only direct fitness), and augmented to allow for the consequences of the actor on others (indirect fitness). Finally, the individual will not usually have full control of its group's fitness, although the group itself might be regarded as having full control of its own fitness [40].

The general point here is that inclusive fitness is not just an accounting method, for mathematically modelling when traits will be favoured by natural selection [14–16,20], it is also the answer to the question of what should organisms appear designed to maximise [10]. No other definition of fitness provides a measure that is both a target of selection and also under the full control of the organism. Hence, natural selection will lead organisms to appear designed to maximise their inclusive fitness, and inclusive fitness is our most general encapsulation of Darwinian fitness [10,41,43]. Our aim here is not to argue that inclusive fitness is the way to answer all evolutionary problems. For example, if you wanted to predict gene dynamics, you would use population genetics. Rather, inclusive fitness is the way to understand organismal design. Similarly, we are not saying that everyone needs to think of organisms as maximizing agents, but rather that you *can* think of them in this way, and that doing so requires inclusive fitness.

It is important to distinguish here between the mathematical modelling and the conceptualisation of natural selection. If the goal is simply to model the evolution of a particular trait, such as altruism, then one can use whatever method is most useful to the problem at hand. However, if the goal is to understand the adaptive rationale for altruism, or any other trait, then the inclusive fitness approach is the only one that fulfils the requisite criteria [5,10,41]. Indeed, because the maths are often more straightforward under the neighbour-modulated fitness approach, a common method is to develop models with neighbour-modulated fitness and then move to inclusive fitness for the purpose of conceptualisation [8,27,28]. Nonetheless, while it can be useful for

theoreticians to mix and match different approaches to their purpose, empirical biologists studying whole organisms need only think about inclusive fitness.

Organisms as Maximizing Agents?

Our third question is why it is useful to have a design principle or maximand? A design principle has been fundamental for linking theoretical and empirical research. When we observe organisms in the field, such as a foraging bird, or an ant tending to her colony's fungus garden, their behaviour has the appearance of design or intention. Inclusive fitness theory provides a link from the gene-frequency dynamics of natural selection to the appearance of design and intention at the individual level [24,41,44]. Specifically, it allows us to conceptualise individuals as trying to maximise something, with that 'something' being inclusive fitness. It is for this reason that inclusive fitness theory has played the central role in the study of adaptation, in fields such as behavioural and evolutionary ecology [3,4]. More generally, by showing how a design principle emerges from the action of natural selection, Darwinism is able to explains the apparent design of the living world [5].

The intuitive advantage of the 'individual as a maximizing agent' analogy is most easily illustrated with a biological example. Consider social insect workers, which forgo the chance to reproduce and instead help to rear their siblings (Figure 3). The inclusive fitness approach suggests that the worker does this because she is closely related to her siblings and so this improves her inclusive fitness. This makes intuitive sense, because the help and relatedness can be easily observed and measured. In contrast, the neighbour-modulated fitness approach suggests that workers help because although the gene carrier is sometimes a worker, who has no neighbour-modulated fitness, at other times the gene carrier is a queen, whose neighbour-modulated fitness is increased owing to help from her workers who also carry copies of the gene. Consequently, on average, individuals carrying genes for worker helping have a higher neighbour-modulated fitness than individuals not carrying such genes. Whilst both descriptions are formally correct, neighbour-modulated fitness does not engage with the intentionality observed in individual behaviour, and hence can be conceptually cumbersome.

The advantage of the maximising agent analogy provided by inclusive fitness theory also becomes obvious when considering how observational and experimental studies are carried out and interpreted. Consider an experiment that manipulates the behaviour of an individual, such as its clutch size decision or level of cooperation. It is standard practice to measure the consequences of that change in behaviour, both for the focal individual and its social partners, such that the inclusive fitness effect of the behaviour
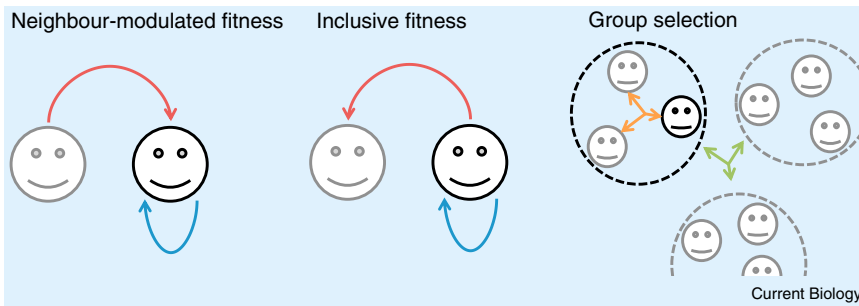
Figure 2. Partitions of natural selection.
The overall correlation between a focal individual's genes and her reproductive success can be decomposed into: (A) the direct effect on the focal individual (blue arrow) and the indirect effect of those genes in the focal individual's social partners (red arrow), (B) the direct effect on the focal individual (blue arrow) and the indirect effect on the focal individual's social partners (red arrow); (C) within-group (orange arrows) and between-group (green arrows) effects. A distinguishing feature of the inclusive fitness approach (B) is that the focal individual has control of the causal fitness effects (the arrows).

can be estimated [3,4]. In contrast, the neighbour-modulated fitness (and group selection) approach emphasises correlations in behaviour between relatives (or group mates) such that, as well as changing the behaviour of the focal individual, the experimenter may also have to change the behaviour of any social partners, by an amount that depended upon their relatedness (or a measure of population structure like F_{ST}). It is not without reason that this isn't the approach usually taken by empirical biologists studying whole organisms.

Another advantage of the maximising agent analogy is that, because it focuses on the perspective of individual actors, it emphasises when design objectives differ, and hence readily identifies evolutionary conflicts of interest. This includes scenarios where there can be some overlap of interest, such as conflicts between family or group members, and between genes within a genome [45–47]. In contrast, the neighbour-modulated fitness approach focuses on the average consequences of a trait across different actors, and so tends to obscure conflicts, providing a more cumbersome way of conceptualising them. It is for this reason that the inclusive fitness approach has led to the identification and study of evolutionary conflicts [3,4,45–47].

Is There Anything Special about Common Ancestry?
We now turn to our fourth question: when considering social adaptations, such as altruism, is there anything special about genetic relatedness that arises from recent common ancestry, as opposed to other assortment mechanisms that could lead to individuals being genetically related? Hamilton [10] was the first to point out that indirect fitness benefits require genetic relatedness for the trait being considered, and not common ancestry (genealogical kinship) *per se*. Hamilton illustrated this by showing how a gene could be favoured if it could identify the presence of copies of itself in other individuals and behave nepotistically towards these other individuals, irrespective of their genealogical relationship to its own carrier. Such assortment mechanisms are usually called 'greenbeards', after Dawkins' [48] thought experiment where the gene causes its carrier to both grow a green beard and also provide help to other individuals with green beards.

The possibility for greenbeard-like effects has been used to argue that assortment is the fundamental issue, and that there is nothing special about genetic relatedness arising from common ancestry (i.e. genealogical kinship) [14,15,20]. In order to address this, we consider the types of adaptation that would arise from genetic relatedness caused by different mechanisms. Greenbeards represent an extreme case, where the assortment mechanism and

social behaviour are encoded on a single gene (or tightly linked genes), and the behaviour is directed towards other carriers of that gene [31]. We would usually expect that assortment mechanisms and the social responses to them are based on complex phenotypes caused by multiple genes scattered across the genome; for example, a gene (or genes) for the cue on which genetic recognition was based, such as smell, a gene for the cue and a gene for the social behaviour [49]. When these genes are scattered, the relatedness between two individuals at the recognition and social behaviour genes will, on average, be that due to common ancestry rather than whether they share the gene for the cue *per se* [49]. This means that natural selection will only favour multiple gene assortment mechanisms if the cue is a reliable indicator of common ancestry [49]. Consequently, we need merely compare the types of adaptation caused by greenbeards versus common ancestry.

Assortment mechanisms, such as greenbeards, are unlikely to lead to significant adaptation because they will be extremely rare and can only lead to traits that are produced by a single or small number of tightly linked genes. First, greenbeards can be outcompeted by cheats that display the beard without also performing the helping behaviour ('falsebeards') [48,50]. Consequently, greenbeards are expected to be rare in the natural world. Second, even if a greenbeard were stable, it would lead to a high genetic relatedness at the greenbeard locus, but not at other parts of the genome (Figure 4A) [51]. This means that any resulting adaptations would have to be constructed by just that greenbeard gene and the genes tightly linked to it. Every other gene in the genome will be united in their attempt to either produce a falsebeard, or to replace the greenbeard with a trait that is responding to genealogical relatedness [52]. Consequently, given that all but the simplest traits are underpinned by multiple genes, greenbeards are unlikely to lead to elaborate adaptation [51,52].

In contrast, a special property of recent common ancestry is that it leads to more or less equal genetic relatedness

Table 2. Design and fitness maximisation.

	Target of selection?	Under own control?
Neighbour-modulated fitness	Yes	No
Inclusive fitness	Yes	Yes
Group fitness	No	Yes/No

In order to represent a quantity that the organism appears designed to maximise, a fitness measure must be both a target of natural selection and also under the organism's sole control. Only inclusive fitness fulfils both of these criteria.

Figure 3. Design and purpose.

If we examine the different castes of a social insect, they appear to have very different designs and purposes. In the termite *Macrotermes bellicosus*, the large queen (A) is essentially an egg laying machine, while workers (B) collect food, maintain and enlarge the nest, and care for the queen, eggs and young, while soldiers defend the nest. Photos by Judith Korb and Volker Salewski.

across the majority of the genome, such that most loci will be pulling in the same direction when constructing adaptations (Figure 4B) [51]. Furthermore, while some methods for generating genetic associations require strong selection differences between alleles, common descent operates regardless of the strength of selection. These points mean that common ancestry can favour the evolution of social adaptations in a way that simply is not possible with other association mechanisms [51].

The empirical data on social traits support the hypothesised special role of common descent, with studies showing that genealogical relatedness matters for a wide range of traits, including sex allocation, policing, conflict resolution, cooperation, altruism, spite, parasite virulence, within group or family conflict, cannibalism, dispersal, alarm calls, eusociality and genetic conflict [3,4,9,50,53]. In contrast, greenbeard effects are incredibly rare [50], and it is still a matter of debate whether the known examples have been selected for because of their greenbeard properties *per se*. More generally, this illustrates that whilst population genetic models provide the gold standard for evolutionary theory, simple one-locus models [14] can be over-interpreted in misleading ways. This is because real-world adaptations are underpinned by multiple loci, distributed across the genome, rather than a single locus [54]. Inclusive fitness is a theory of whole-organism biology.

Genes Versus Individuals?

The idea that we can treat individuals as maximising agents is based on the implicit assumption that we can effectively ignore the consequences of genetic conflict within individuals. However, natural selection is driven by changes in gene frequencies, and 'selfish genes' can be favoured to increase their own transmission, even if this incurs a cost at the level of the individual [47]. So, why do we not give up on the idea of individuals as maximising agents, and just conceptualise natural selection via gene dynamics?

One reason is that, to explain adaptation, we are almost forced to think in terms of a maximand and at the individual level. Even at the level of the gene, we would still want to know what the maximand is, and the answer is 'the inclusive fitness of the gene' [54]. But, adaptation is manifest at the level of the individual, and results from the cooperation of multiple genes distributed across the genome, not just the operation of a single gene [54]. Consequently, we require a theory that explains how and why genes pull together at the individual level, and the form of adaptation that this will lead to. Inclusive fitness theory does this. Indeed, as we saw in the previous section, population genetic models can be misinterpreted in terms of the adaptations we would expect at the individual level.

Another reason is that the individual approach provides a heuristic that facilitates the interplay between theory and data. By black boxing the genes, we can focus on other aspects of biology, such as ecology and behaviour. This makes it easier to both develop models, and test the robustness of those models to changes in the underlying biological parameters [8,55,56]. A focus on individuals also makes it easier to spot when there will be conflicts between individuals, such as between parents and their offspring, or between siblings [46]. Furthermore, working out the exact solution to a more specific genetic model might not be very helpful, because we will rarely know the underlying genetic architecture of a trait. The individual approach assumes that genetic architecture does not constrain adaptation (phenotypic gambit), and hence provides a robust approximation to a wide range of genetic models that can be applied broadly across a range of species [21,32].

A general point here is that the individual approach trades off generality and applicability versus exactness, in a way that facilitates progress. Furthermore, it does so when the exactness would usually have to be based on guesswork. Nonetheless, we stress the importance of seeing this as a heuristic approach, which should only be used when the advantages outweigh the disadvantages, which is best judged empirically. The progress made with the individual approach, in fields such as behavioural and evolutionary ecology, suggest that the benefits usually far outweigh the costs [3,4,56].

The disadvantages of ignoring genetic conflict will often be minor, because we expect it to have relatively little influence on adaptation at the individual level. First, for most traits, such as foraging or predator avoidance, the only way that a gene can increase its transmission to the next generation is by increasing the fitness of the individual carrying it, so
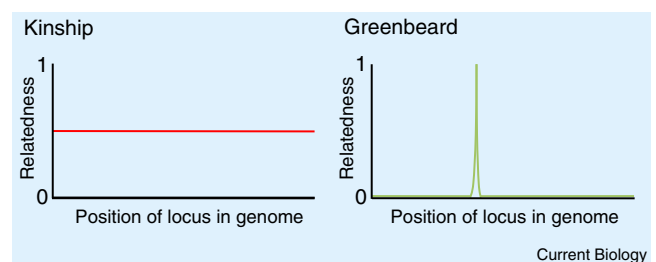

Figure 4. Relatedness through common ancestry and assortment.

(A) Common ancestry: the genetic relatedness between outbred diploid siblings is the same at all autosomal loci. (B) If genetic similarity is caused only by a single greenbeard gene, then relatedness is expected to drop off on either side of the greenbeard. The rate of drop off will depend upon recombination rates, population structuring, etc. Adapted from [51].
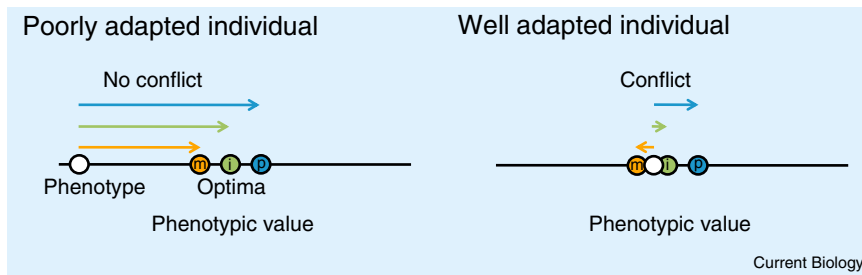
Figure 5. Conflict and fitness maximization.
The figure shows a scenario where there is a conflict of interest over the optimum phenotype between genes with maternal origin (m), paternal origin (p), or which have no information about their origin (i) [45]. The latter also represents the perspective of a gene trying to maximise the inclusive fitness of the individual. (A) If organisms were poorly fitted to their environments, all their genes would be pulling in much the same direction, towards a distant set of closely-coinciding optima. (B) It is only upon closing in upon the individual optimum that minor quantitative disagreements between genes will come into play.

there will be no conflict. Consequently, we would expect genetic conflict to be limited to traits where transmission can be altered, such as the offspring sex ratio [47]. Second, when a gene increases its own transmission to the detriment of the other genes in the genome, all of those other genes will be united in what Leigh [57] termed the 'parliament of the genes' to suppress the selfish gene. So, for example, while selfish sex ratio distorters can be common, their presence is strongly correlated with suppressors, such that they have little influence on the population sex ratio [58,59]. Third, complex adaptations are underpinned by multiple loci, and so we would not expect selfish genes acting alone to be able to manifest complex adaptation [54]. Fourth, even in these cases where we expect conflict, the overall effect of genes at multiple loci pulling in different directions will often cancel out at the individual level. The very fact that we can observe such conflicts suggests that the organism has attained near perfection in its adaptation (Figure 5). Furthermore, one of the best ways to discover genetic conflicts empirically is to detect departures from inclusive fitness maximisation at the individual level. All of these points clarify the utility of the individual as maximising agent analogy in the study of adaptation.

Asking the Wrong Questions

The above points make it clear that, in order to challenge the inclusive fitness paradigm, it would be necessary to show that there is a more useful maximand (design objective) than inclusive fitness. None of the critiques of inclusive fitness have even suggested an alternative maximand, let alone compared their relative utility [12–20]. These critiques have missed that the major purpose of inclusive fitness is to provide an answer to the question of what organisms should appear designed to maximise. Another possibility could have been to show that the individual as a maximizing agent analogy is not useful for linking theory to data, in conflict with evidence from fields such as behavioural and evolutionary ecology [3,4,9], or that there is a more useful approach. Again, none of the critiques even touched upon this. Indeed, the critiques themselves used intentional language such as 'altruism', the use of which is formally justified by the maximizing agent analogy [60], and so they appear to endorse rather than criticize this approach.

Instead of engaging with such scientifically-relevant issues, the papers criticizing inclusive fitness theory simply reinvented old problems that were solved decades ago, and are hence largely irrelevant [21,22]. For example, arguing that inclusive fitness requires restrictive assumptions, that it is less general, or that it does not make predictions that are different from those yielded by other methods, such as population genetics [12–20]. It has long been known that inclusive fitness theory, and all of the other partitions we have discussed here, can be used to make either relatively general models with few assumptions, for the purpose of providing a conceptual overview, or more specific models with more assumptions, when that is useful for predicting the evolution of specific traits [8,27,28,33,61]. To give another example, inclusive fitness theory was not developed as a competitor to the theory of natural selection [15]. Instead, it is the answer to the question of what is the design principle (maximand) that emerges from the process of natural selection, and so we expect that, when the same assumptions are made, different methods should yield the same prediction.

It is not clear why the role of inclusive fitness as a maximand has been ignored in the recent critiques of inclusive fitness theory. A reading of even just the abstract of Hamilton's [10] original inclusive fitness paper makes clear that his motivation was to find a maximand of natural selection, and not just an accounting method: 'a quantity is found which incorporates the maximizing property of Darwinian fitness. This quantity is named 'inclusive fitness'. Species following the model should tend to evolve behaviour such that each organism appears to be attempting to maximize its inclusive fitness'. Since then, the concept of inclusive fitness maximisation has been central to fields such as behavioural and evolutionary ecology [3,4,32,43,62].

Another criticism of inclusive fitness theory has been that it focuses on common ancestry, which is just one mechanism for producing genetic associations [12,14]. This criticism is misguided on two counts. First, the possibility for other association mechanisms was already clarified almost 40 years ago, within the framework of inclusive fitness theory [10,31]. Second, the aim of inclusive fitness theory is not just to explain when a single gene will spread, but rather to explain complex organismal adaptation. Adaptations are underpinned by multiple genes distributed across the genome, and so we need a mechanism that leads to genes pulling in the same direction when constructing social adaptations. Recent common ancestry neatly solves this problem [51], and has been central to explaining the empirical data on all forms of social trait, from cannibalism to altruism [3,4]. None of the critiques of inclusive fitness theory have suggested an equally valid or better way to unite the interest of genes.

By asking the wrong questions, and reinventing long-solved problems, these critiques have removed attention from more biologically interesting questions on the use of inclusive fitness theory. For example, in bacteria, mobile

genetic units, such as plasmids, can move horizontally, leading to the genes on the plasmid having a different inclusive fitness optimum [63]. When is this optimum sufficiently different from that of the chromosomes that we have to think of the cell as fragmented into two maximizing agents (chromosomes versus plasmid), rather than a single maximising agent? Other unresolved issues include the impact on inclusive fitness maximisation of non-additive fitness effects [43,54], or traits not being conditionally expressed according to the actor class, such as queen or worker [8].

Conclusions

It is useful to consider what we would require from a theory of adaptation, if we were to develop one from scratch. We would require the theory to do three things: first, capture the quantitative dynamics of natural selection, so that it can predict the evolution of specific traits; second, provide a maximand or design objective for organismal traits; third, explain why genes across the genome pull in the same direction, when constructing complex adaptations. The beauty and unique position of inclusive fitness theory is that it fulfils all three of these requirements. None of the critiques of inclusive fitness theory have suggested an alternative theory of adaptation that fulfils these criteria.

Supplemental Information

Supplemental information providing conceptual background can be found with this article online at http://dx.doi.org/10.1016/j.cub.2013.05.031.

References

1. Paley, W. (1802). Natural Theology (London: Wilks & Taylor).
2. Darwin, C. (1859). On the Origin of Species by Means of Natural Selection, or, the Preservation of Favoured Races in the Struggle for Life (London, UK: John Murray).
3. Davies, N.B., Krebs, J.R., and West, S.A. (2012). An Introduction to Behavioural Ecology, Fourth Edition (Oxford, England: Blackwell Scientific Publications).
4. Westneat, D.F., and Fox, C.W. (2010). Evolutionary Behavioral Ecology (Oxford: Oxford University Press).
5. Gardner, A. (2009). Adaptation as organism design. Biol. Lett. 5, 861–864.
6. Alcock, J. (2009). Animal Behavior (Sunderland, Massachusetts: Sinauer Associates).
7. Manning, A., and Dawkins, M.S. (2012). An Introduction to Animal Behavioiur, Sixth Edition (Cambridge University Press).
8. Frank, S.A. (1998). Foundations of Social Evolution (Princeton: Princeton University Press).
9. Bourke, A.F.G. (2011). Principles of Social Evolution (Oxford: Oxford University Press).
10. Hamilton, W.D. (1964). The genetical evolution of social behaviour, I & II. J. Theor. Biol. 7, 1–52.
11. Hamilton, W.D. (1970). Selfish and spiteful behaviour in an evolutionary model. Nature 228, 1218–1220.
12. Wilson, E.O., and Hölldobler, B. (2005). Eusociality: origin and consequences. Proc. Natl. Acad. Sci. USA 102, 13367–13371.
13. Wilson, D.S., and Wilson, E.O. (2007). Rethinking the theoretical foundation of sociobiology. Q. Rev. Biol., 327–348.
14. Fletcher, J.A., and Doebeli, M. (2009). A simple and general explanation for the evolution of altruism. Proc. Roy. Soc. Lond. B 276, 13–19.
15. Nowak, M.A., Tarnita, C.E., and Wilson, E.O. (2010). The evolution of eusociality. Nature 466, 1057–1062.
16. Doebeli, M. (2010). Inclusive fitness is just bookkeeping. Nature 467, 661.
17. van Veelen, M. (2009). Group selection, kin selection, altruism and cooperation: when inclusive fitness is right and when it can be wrong. J. Theoret. Biol. 259, 589–600.
18. van Veelen, M., Garcia, J., Sabelis, M.W., and Egas, M. (2012). Group selection and inclusive fitness are not equivalent; the Price equation versus models and statistics. J. Theoret. Biol. 299, 64–80.
19. Traulsen, A. (2010). Mathematics of kin- and group-selection: formally equivalent? Evolution 64, 316–323.
20. Damore, J.A., and Gore, J. (2012). Understanding microbial coooperation. J. Theoret. Biol. 299, 31–41.
21. Gardner, A., West, S.A., and Wild, G. (2011). The genetical theory of kin selection. J. Evol. Biol. 24, 1020–1043.
22. Rousset, F., and Lion, S. (2011). Much ado about nothing: Nowak et al.'s charge against inclusive fitness theory. J. Evol. Biol. 24, 1386–1392.
23. Bourke, A.F.G. (2011). The validity and value of inclusive fitness theory. Proc. Roy. Soc. Lond. B 278, 3313–3320.
24. Fisher, R.A. (1930). The Genetical Theory of Natural Selection (Oxford: Clarendon).
25. Price, G.R. (1970). Selection and covariance. Nature 227, 520–521.
26. Maynard Smith, J. (1983). Models of evolution. Proc. Roy. Soc. Lond. B 219, 315–325.
27. Taylor, P.D., and Frank, S.A. (1996). How to make a kin selection model. J. Theor. Biol. 180, 27–37.
28. Rousset, F. (2004). Genetic Structure and Selection in Subdivided Populations (Princeton, NJ: Princeton University Press).
29. Moore, A.J., Brodie, E.D.I., and Wolf, J.B. (1997). Interacting phenotypes and the evolutionary process: I. Direct and indirect genetic effects of social interactions. Evolution 51, 1352–1362.
30. Price, G.R. (1972). Extension of covariance selection mathematics. Ann. Hum. Genet. 35, 485–490.
31. Hamilton, W.D. (1975). Innate social aptitudes of man: an approach from evolutionary genetics. In Biosocial Anthropology, R. Fox, ed. (New York: Wiley), pp. 133–155.
32. Grafen, A. (1984). Natural selection, kin selection and group selection. In Behavioural Ecology: An Evolutionary Approach, 2nd Edition, J.R. Krebs and N.B. Davies, eds. (Oxford, UK: Blackwell Scientific Publications), pp. 62–84.
33. Frank, S.A. (1986). Hierarchical selection theory and sex ratios. I. General solutions for structured populations. Theoret. Popul. Biol. 29, 312–342.
34. Queller, D.C. (1992). Quantitative genetics, inclusive fitness, and group selection. Am. Nat. 139, 540–558.
35. Taylor, P.D., Wild, G., and Gardner, A. (2007). Direct fitness or inclusive fitness: how shall we model kin selection. J. Evol. Biol. 20, 301–309.
36. Lion, S., Jansen, V.A.A., and Day, T. (2011). Evolution in structured populations: beyond the kin versus group debate. Trends Ecol. Evol. 26, 193–201.
37. Nee, S., West, S.A., and Read, A.F. (2002). Inbreeding and parasite sex ratios. Proc. Roy. Soc. Lond. B 269, 755–760.
38. Frank, S.A. (2012). Natural Selection. III. Selection versus transmission and the levels of selection. J. Evol. Biol. 25, 227–243.
39. West, S.A., Griffin, A.S., and Gardner, A. (2008). Social semantics: how useful has group selection been? J. Evol. Biol. 21, 374–385.
40. Gardner, A., and Grafen, A. (2009). Capturing the superorganism: a formal theory of group adaptation. J. Evol. Biol. 22, 659–671.
41. Grafen, A. (2007). The formal Darwinism project: a mid-term report. J. Evol. Biol. 20, 1243–1254.
42. Grafen, A. (2002). A first formal link between the Price equation and an optimization program. J. Theor. Biol. 217, 75–91.
43. Grafen, A. (2006). Optimisation of inclusive fitness. J. Theoret. Biol. 238, 541–563.
44. Dawkins, R. (1978). Replicator selection and the extended phenotype. Zeitschrift für Tierpsychologie 47, 61–76.
45. Haig, D. (2002). Genomic Imprinting and Kinship (New Brunswick, NJ: Rutgers University).
46. Trivers, R.L. (1974). Parent-offspring conflict. Am. Zool. 14, 249–264.
47. Burt, A., and Trivers, R. (2006). Genes in Conflict: The Biology of Selfish genetic Elements (Cambridge, Massachusetts: Harvard University Press).
48. Dawkins, R. (1976). The Selfish Gene (Oxford: Oxford University Press).
49. Grafen, A. (1990). Do animals really recognise kin? Anim. Behav. 39, 42–54.
50. Gardner, A., and West, S.A. (2010). Greenbeards. Evolution 64, 25–38.
51. Grafen, A. (1985). A geometric view of relatedness. Oxford Surv. Evol. Biol. 2, 28–89.
52. Biernaskie, J.M., West, S.A., and Gardner, A. (2011). Are greenbeards intragenomic outlaws? Evolution 65, 2729–2742.
53. West, S.A. (2009). Sex Allocation (Princeton: Princeton University Press).
54. Gardner, A., and Welch, J.J. (2011). A formal theory of the selfish gene. J. Evol. Biol. 24, 1801–1813.
55. Pen, I., and Weissing, F.J. (2000). Sexual selection and the sex ratio: an ESS analysis. Selection 1, 59–69.
56. Parker, G.A., and Maynard Smith, J. (1990). Optimality theory in evolutionary biology. Nature 348, 27–33.

57. Leigh, E.G. (1971). Adaptation and Diversity (San Francisco: Freeman, Cooper and Company).

58. Atlan, A., Mercot, H., Landre, C., and Montchamp-Moreau, C. (1997). The sex-ratio trait in *Drosophila simulans*: geographical distribution of distortion and resistance. Evolution *51*, 1886–1895.

59. Taylor, D.R. (1999). Genetics of sex ratio variation among natural populations of a dioecious plant. Evolution *53*, 55–62.

60. Grafen, A. (1999). Formal Darwinism, the individual-as-maximizing-agent analogy and bet-hedging. Proc. Roy. Soc. Lond. B *266*, 799–803.

61. Queller, D.C. (1992). A general model for kin selection. Evolution *46*, 376–380.

62. Grafen, A. (1991). Modelling in behavioural ecology. In Behavioural Ecology, an Evolutionary Approach, J.R. Krebs and N.B. Davies, eds. (Oxford: Blackwell), pp. 5–31.

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

64. Maynard Smith, J. (1969). The status of Neo-Darwinism. In Towards a Theoretical Biology 2. Sketches, C.H. Waddington, ed. (Edinburgh: Edinburgh University Press), pp. 82–94.

Supplemental information

Adaptation and inclusive fitness

Stuart A. West and Andy Gardner

The dynamics of natural selection

Here, we more formally address the first question of our main text, concerning the equivalence and generality of the alternative modeling approaches that are used to study adaptive evolution. The major contenders all stem from the same basic principle, that natural selection drives genetic change of biological populations. Darwin's [S1] original argument was that natural selection will favour those heritable traits that are associated with greater reproductive success, leading individual organisms to become better fitted to their environments. Fisher [S2] later made these ideas mathematically concrete, by unifying Darwinian evolution with Mendelian genetics. Importantly, Fisher formally defined Darwinian fitness as the individual's asymptotic contribution of genes to future generations, or 'reproductive value'. This genetical theory of natural selection was elegantly encapsulated by Price [S3], whose equation states that the change in the population average of any heritable trait $\Delta E(g)$ is equal to the covariance of that trait g with relative fitness v:

$$\Delta_S E(g) = cov(v,g). \tag{1}$$

This clarifies that the key criterion for natural selection to favour any trait is that genes for this trait are positive correlated with individual fitness. However, correlation may obscure causation, particularly when considering the evolution of altruistic traits that appear, at first sight, to be precluded by equation (1). Consequently, social evolution researchers have found it helpful to split equation (1) into more meaningful, causal relations. Debate has focused on the relative usefulness of three basic approaches.

First, the "personal fitness" or "neighbour modulated fitness" approach decomposes the overall correlation between an individual's genes and her fitness into a direct and an indirect effect [S4-S9]. The direct effect owes to the expression of those genes in the individual's own phenotype. The indirect (kin selection) effect, that owes to the expression of copies of those same genes in the phenotypes of the individual's social partners, and contributes to the action of selection in proportion to the genetic relatedness between the individual and her social partners (Figure 2a). Mathematically, this may be written as:

$$\Delta_S E(g) = \beta(v,g|g')cov(g,g) + \beta(v,g'|g)cov(g',g), \tag{2}$$

where g' denotes the genetic value of the individual's social partner [S4,S10-S12]. The direct fitness effect is negative for altruistic behaviours (costly), and hence is often *denoted* $\beta(v,g|g') = -c$. The indirect fitness effect is positive for altruistic behaviours (beneficial), and is denoted $\beta(v,g'|g) = b$. The ratio $cov(g',g)/cov(g,g) = r$ defines the genetic relatedness of social partners [S4,S11-S14]. Consequently, the condition for natural selection to favour altruism (or, indeed, any trait) is $-c + br > 0$,

which is Hamilton's rule of kin selection [S8,S15]. The theory of indirect genetic effects (IGEs) takes an essentially similar approach, with more explicit focus upon the phenotypic effects, rather than just the fitness effects, of those genes [S16].

Second, the "inclusive fitness" approach uses the same basic partition of natural selection into direct and indirect effects, but conceptualizes the indirect effect in a different ways [S8]. In particular, whilst the neighbour modulated fitness approach considers the impact $\beta(v,g'|g)$ of the genes g' of social partners on the focal individual's own fitness v, the inclusive fitness approach instead considers the impact $\beta(v',g|g')$ of the focal individual's genes g on the fitness v' of social partners. Mathematically, this may be written as:

$$\Delta_S E(g) = \beta(v,g|g')cov(g,g) + \beta(v',g|g')cov(g',g), \qquad (3)$$

This difference in assigning fitness effects makes a difference at the level of the individual, because an individual's inclusive fitness need not be equal to her neighbour modulated fitness. For example, a worker ant is able to achieve inclusive fitness despite being sterile. Here, some of the reproductive success of her queen, that is assigned to the queen in the context of the neighbour modulated-fitness approach to kin selection, is reassigned to the worker in the context of the inclusive-fitness approach to kin selection. However, in aggregate, over the whole population, the two ways of computing indirect effects always add up to the same thing ($\beta(v,g'|g)$ = $\beta(v',g|g')$ = b). Hence, both approaches to kin selection yield the same criterion for selection to favour any trait: $-c + br > 0$.

The two approaches lead to a different conceptualization of why relatedness matters. The inclusive fitness approach emphasizes that individuals value the reproduction of relatives because they share genes. That is, the genetic relatedness acts as a currency converter describing the valuation that an individual places upon the reproduction of her relatives. In contrast, the neighbour modulated fitness approach emphasizes that the heritable traits of social partners are more likely to be similar when they are more closely related. That is, to the extent that an individual interacts with relatives, her social partners will tend to mirror any tendency that she has for exhibiting a trait.

Third, the group selection approach decomposes the overall correlation between an individual's genes and her fitness into within- and between-group effects. The within-group effect describes the average association between genes and fitness within social groups in the population, and contributes to the action of selection in proportion to the heritable variation that exists within social groups. The between-group effect, that describes the association between genes and fitness at the group level, and contributes to the action of selection in proportion to the heritable variation that exists between social groups (Figure 2b). Mathematically, this may be written as:

$$\Delta_S E(g) = E_I(cov_J(v_{ij},g_{ij})) + cov_G(v_i,g_i), \qquad (4)$$

where each group is assigned an index i∈I and each individual in a group is assigned an index j∈J. The within-group effect may be negative for some altruistic behaviours, and the between-group effect may be positive for some altruistic behaviours. However, altruism that aids close kin to the detriment of unrelated individuals within the individual's social group may lead to a positive within-group effect and a negative

between-group effect, so the group selection partition cannot be used to define altruism (contra [S17]).

The above results have two implications. First, these different methods do not represent competing hypotheses about how social evolution occurs. Instead, they are just different mathematical ways of dividing up the dynamics of natural selection, in terms of how genes influence their transmission to future generations [S4-S8,S18-S21]. Second, none of these different methods require restrictive assumptions such as pairwise interactions or linear fitness effects [S4,S7,S11,S12]. Restrictive assumptions are sometimes made when modeling specific scenarios, but this is true for all approaches, and is only done by choice (not requirement), when it is both appropriate and useful for making testable predictions [S4,S5,S7,S11,S13].

It is useful to distinguish between our aims and methods in this section, with that used when making testable predictions for specific traits. We have used the Price equation approach as a 'meta-model' to provide a general overview and show how the different approaches are related [S11,S22]. In contrast, when making specific predictions for specific traits it is usually more appropriate to combine methods such as game theory with more explicit biological assumptions [S4,S5,S7,S13]. Furthermore, when making more testable models, the aim is usually to predict relationships between specific biological parameters, such as how a trait should vary with some environmental condition, rather than starting with or focusing on Hamilton's rule. For example, how the offspring sex ratio should vary with the number of female parasitic wasps on a patch, or, in social insects, the number of times that the queen has mated [S23-S25].

Alternative Partitions

Hamilton [S8] developed both the neighbour modulated fitness and the inclusive fitness approaches to kin selection, and then linked them to group selection methodology [S18] that had been developed by Price [S26]. The neighbour modulated fitness method was initially "rather unwieldy" [S8] and so it was "much easier to use inclusive fitness" [S9]. Taylor and Frank [S4-S,S27-S29] showed how the standard differentiation methods of optimization theory could be applied to neighbour modulated fitness when relatives interact. This revolutionized social evolution theory, and has since been developed further by Rousset and others, making it easier to develop more general models, which are led by the biology, and can then be interpreted with inclusive fitness [S4-S7,S27-S31]. Consequently, this interplay between neighbour modulated fitness and inclusive fitness has become the standard in most areas of social evolution theory where the aim is to develop empirically testable models. Furthermore, it has even been shown how this enables solutions to problems that had not previously been possible [S30]. Taylor and Frank [S5] termed their approach the 'direct fitness' method, but we refer to neighbour modulated fitness or personal fitness, because: (a) these terms had already been used to describe these methods [S8,S9]; (b) we wish to avoid confusion with the direct component of neighbour modulated fitness or inclusive fitness [S32].

We suggest that confusion has arisen in this area for two reasons. First, debate has centered around specific traits such as altruism or eusociality, rather than how organisms are designed (i.e. what is our most general solution to the problem of adaptation). Second, an abstract theoretical literature has emerged, that is almost

totally divorced from empirical science. Indeed, it is notable that in the areas of evolutionary and behavioural ecology, where theoretical and empirical work go hand in hand, the debate over the usefulness of inclusive fitness theory and the idea of individuals as maximizing agents was resolved decades ago [S20,S33-S36].

When comparing the neighbour modulated and inclusive fitness approaches, an analogy can be made with the distinction between the active and the passive voice. The active voice / inclusive fitness interpretation makes the actor the subject, whereas the passive voice / neighbour modulated fitness interpretation makes the recipient the subject. An example of active versus passive voice is: "I performed the experiment" versus "the experiment was performed". An example of inclusive versus neighbour modulated fitness is: "the individual increases her inclusive fitness" versus "the neighbour modulated fitness of the individual is, on average, greater if she carries the gene". The active voice is generally preferred, because it enlivens the sentence and makes it less dry and dull. Likewise, the inclusive fitness approach enlivens biology by rendering it in terms of agents and their agendas. This is a matter of human psychology, but it makes sense for scientific discourse to work with, rather than against, our psychological make up.

Another way of looking at this is that the inclusive fitness approach clarifies from the focus of intentionality in individual behaviour, whereas the neighbour modulated fitness approach clarifies from the gene's eye view. The use of intentional language, such as 'altruism', by evolutionary biologists is justified when it can be linked to an individual acting as a maximizing agent analogy, in terms of its fitness [S37]. The authors of the papers which have criticized the inclusive fitness approach, still use such intentional language, and by implication, the individual as a maximising agent analogy, even though they are critising the framework upon which formally justifies it [S17,S38-S45]. One of our referees suggested that this apparent contradiction could be explained by the fact that talking about altruism helps publication in certain journals!

Fitness Maximization

In the main text we have distinguished between whether a technique provides an accounting method or a maximand. It is sometimes assumed that inclusive fitness was developed purely as an accounting method. However, Hamilton [S8] made clear that a major motivation was to find a maximand of natural selection:
> "a quantity is found which incorporates the maximizing property of Darwinian fitness. This quantity is named "inclusive fitness". Species following the model should tend to evolve behaviour such that each organism appears to be attempting to maximize its inclusive fitness." (abstract, page 1)
> "it is clear that the quantity that will tend to maximize, if any, is R*, the mean inclusive fitness" (page 7)
> "Here then we have discovered a quantity, inclusive fitness, which under the conditions of the model tends to maximize in much the same way that fitness tends to maximize in the simpler classical model." (page 8)
> This search for a maximand was clearly why Hamilton went on to develop inclusive fitness, when he had already developed neighbour modulated fitness as an accounting method for kin selection [S8]. As Dawkins [S46] put it: "Inclusive fitness

4

may be defined as what property of an individual organism which will appear to be maximized when what is really being maximized is gene survival" . Technical and non-technical overviews of inclusive fitness maximization are provided elsewhere [S35,S47,S48].

Moreover, we emphasize that inclusive fitness provides a general result for what natural selection leads organisms to appear designed to maximise. It is often assumed that inclusive fitness is only relevant when relatives interact. However, a very special case of inclusive fitness theory (or Hamilton's rule) is when r=0, showing that when there are no indirect effects, individuals can be expected to maximise their own reproductive success [S35,S49]. That is, inclusive fitness theory clarifies the precise conditions under which the classical Darwinian result of neighbour modulated fitness maximization obtains.

Inclusive fitness and empirical studies

An introduction to the relevant empirical literature is provided elsewhere. A key point here is that the purpose of empirical studies is not to measure inclusive fitness (contra [S40]), but rather to test the predictions made by inclusive fitness theory [S20,S35,S50]. One area where more complicated manipulations would be required to empirically detect them is when traits have synergistic effects [S51-S54].

Conflict between maximizing agents

Another advantage of the maximizing agent analogy is that, because it focuses on the perspective of individual actors, it emphasizes when design objectives differ, and hence readily identifies evolutionary conflicts of interest. This includes conflicts between family or group members, who may also share some overlap in their interests. To give an example, consider the conflict of interest between mother and offspring in the allocation of maternal resources among the offspring, with diminishing returns on investment into offspring fitness. Taking an inclusive fitness perspective, we can say that an offspring is more related to herself than she is to her siblings, and so she would prefer that maternal resources be preferentially allocated to herself than to her siblings, whilst her mother is equally related to all of her offspring and so would prefer a more equitable allocation of resources over her brood. This neatly describes the tension between offspring and maternal traits in a single sentence.

Taking a neighbour modulated fitness approach, we would say that an offspring who carries a rare gene that, when carried by an offspring, leads to a greater allocation of maternal resources to that offspring than to her siblings, will enjoy an increase in fitness owing to her obtaining more than her fair share of maternal resources, but will also suffer a decrease in fitness owing to her siblings behaving similarly selfishly and, all else being equal, the increment is twice the decrement because, whilst she definitely carries a copy of the gene, each sibling only carries the gene with probability one half, and so, all else being equal, her fitness is increased if she carries the gene, and hence offspring are favoured to extract more than their fair share of maternal resources. In contrast, an offspring who carries a rare gene that, when carried by a mother, leads to an unequal allocation of maternal resources to one of her offspring, will enjoy an increase in fitness if she happens to be the lucky offspring but will suffer a decrease in fitness if she is one of the unlucky offspring and, since she is

just as likely to share this gene in common with her mother if she is lucky or unlucky, all else being equal, her fitness is not increased by carrying the gene (and, indeed, is reduced if inequitable distribution of maternal resources reduces the overall fitness of the brood), and hence mothers are not favoured to distribute their resources unevenly among their offspring.

The above example demonstrates that the inclusive fitness approach is a less cumbersome way of conceptualising conflicts of interest. Furthermore, with more complicated scenarios, where populations are structured into classes such as sex or caste, the neighbour modulated fitness approach becomes increasingly unwieldy. For example, if offspring could be male or female, then we would have to add in another class of individual to average over. Inclusive fitness deals with this far more simply, by varying how different classes of individuals are valued, according to their relatedness and reproductive value.

Greenbeards and adaptation

In the main text, we pointed out that because all but the simplest traits are produced by multiple genes, greenbeards are unlikely to be able to give rise to complex adaptation. An interesting possible exception to this, which requires further investigation, are the genetic units which move horizontally in microbes, such as plasmids. These units can be of non-trivial size, relative to the rest of the bacterial genome, and it is notable that possibly the most common form of greenbeard - bacteriocin warfare - is often underpinned by plasmid-born genes [S55].

The extent to which this leads to the various adaptations of the bacterial cell having different design objectives is not clear [S56]. For example, would population structuring or conflict over plasmid transmission/expression tend to align the interests of plasmids with the rest of the bacterial genome? Or do we need to consider these different genetic entities as separate inclusive fitness maximizing agents? Both theoretical and empirical work is required to address such questions [S57-S61].

Supplemental references

S1.   Darwin, C. (1859). On the Origin of Species by Means of Natural Selection, or, the Preservation of Favoured Races in the Struggle for Life, (London, UK: John Murray).

S2.   Fisher, R.A. (1930). The Genetical Theory of Natural Selection, (Oxford: Clarendon).

S3.   Price, G.R. (1970). Selection and covariance. Nature *227*, 520-521.

S4.   Frank, S.A. (1998). Foundations of Social Evolution, (Princeton: Princeton University Press).

S5.   Taylor, P.D., and Frank, S.A. (1996). How to make a kin selection model. J. Theor. Biol. *180*, 27-37.

S6.   Taylor, P.D., Wild, G., and Gardner, A. (2007). Direct fitness or inclusive fitness: how shall we model kin selection. Journal of Evolutionary Biology *20*, 301-309.

S7.   Rousset, F. (2004). Genetic structure and selection in subdivided populations, (Princeton, NJ: Princeton University Press).

S8. Hamilton, W.D. (1964). The genetical evolution of social behaviour, I & II. J. Theor. Biol. *7*, 1-52.

S9. Maynard Smith, J. (1983). Models of evolution. Proc. Roy. Soc. Lond. B *219*, 315-325.

S10. Hamilton, W.D. (1970). Selfish and spiteful behaviour in an evolutionary model. Nature *228*, 1218-1220.

S11. Gardner, A., West, S.A., and Wild, G. (2011). The genetical theory of kin selection. Journal of Evolutionary Biology *24*, 1020-1043.

S12. Queller, D.C. (1992). A general model for kin selection. Evolution *46*, 376-380.

S13. Frank, S.A. (1983). Theoretical and empirical studies of sex ratios, mainly in fig wasps. (University of Florida).

S14. Orlove, M.J., and Wood, C.L. (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.

S15. Hamilton, W.D. (1963). The evolution of altruistic behaviour. Am. Nat. *97*, 354-356.

S16. Moore, A.J., Brodie, E.D.I., and Wolf, J.B. (1997). Interacting phenotypes and the evolutionary process: I. Direct and indirect genetic effects of social interactions. Evolution *51*, 1352-1362.

S17. Wilson, D.S., and Wilson, E.O. (2007). Rethinking the theoretical foundation of sociobiology. Q. Rev. Biol., 327-348.

S18. Hamilton, W.D. (1975). Innate social aptitudes of man: an approach from evolutionary genetics. In Biosocial Anthropology, R. Fox, ed. (New York: Wiley), pp. 133-155.

S19. Frank, S.A. (1986). Hierarchical selection theory and sex ratios. I. General solutions for structured populations. Theoret. Popul. Biol. *29*, 312-342.

S20. Grafen, A. (1984). Natural selection, kin selection and group selection. In Behavioural Ecology: An Evolutionary Approach, 2nd Edition, J.R. Krebs and N.B. Davies, eds. (Oxford, UK: Blackwell Scientific Publications), pp. 62-84.

S21. Queller, D.C. (1992). Quantitative genetics, inclusive fitness, and group selection. American Naturalist *139*, 540-558.

S22. Frank, S.A. (2012). Natural Selection. IV. The Price equation. Journal of Evolutionary Biology *25*, 1002-1019.

S23. Boomsma, J.J., and Grafen, A. (1991). Colony-level sex ratio selection in the eusocial Hymenoptera. J. Evol. Biol. *4*, 383-407.

S24. Hamilton, W.D. (1967). Extraordinary sex ratios. Science *156*, 477-488.

S25. West, S.A. (2009). Sex Allocation, (Princeton: Princeton University Press).

S26. Price, G.R. (1972). Extension of covariance selection mathematics. Annals of Human Genetics *35*, 485-490.

S27. Frank, S.A. (1995). Mutual policing and repression of competition in the evolution of cooperative groups. Nature *377*, 520-522.

S28. Frank, S.A. (1996). Models of parasite virulence. Q. Rev. Biol. *71*, 37-78.

S29. Frank, S.A. (1997). Multivariate analysis of correlated selection and kin selection, with an ESS maximization method. J. Theor. Biol. *189*, 307-316.

S30. Nee, S., West, S.A., and Read, A.F. (2002). Inbreeding and parasite sex ratios. Proc. Roy. Soc. Lond. B *269*, 755-760.

S31.   Wild, G. (2011). Direct fitness for dynamic kin selection. Journal of Evolutionary Biology *24*, 1020-1043.

S32.   West, S.A., Griffin, A.S., and Gardner, A. (2008). Social semantics: how useful has group selection been? Journal of Evolutionary Biology *21*, 374-385.

S33.   Davies, N.B., Krebs, J.R., and West, S.A. (2012). An Introduction to Behavioural Ecology, Fourth Edition, Fourth Edition, (Oxford, England: Blackwell Scientific Publications).

S34.   Grafen, A. (1990). Biological signals as handicaps. Journal of Theoretical Biology *144*, 517-546.

S35.   Grafen, A. (1991). Modelling in behavioural ecology. In Behavioural Ecology, an Evolutionary Approach, J.R. Krebs and N.B. Davies, eds. (Oxford: Blackwell), pp. 5-31.

S36.   Westneat, D.F., and Fox, C.W. (2010). Evolutionary Behavioral Ecology, (Oxford: Oxford University Press).

S37.   Grafen, A. (1999). Formal Darwinism, the individual-as-maximizing-agent analogy and bet-hedging. Proc. Roy. Soc. Lond. B *266*, 799-803.

S38.   Wilson, E.O., and Hölldobler, B. (2005). Eusociality: origin and consequences. Proc. Natl. Acad. Sci. U.S.A. *102*, 13367-13371.

S39.   Fletcher, J.A., and Doebeli, M. (2009). A simple and general explanation for the evolution of altruism. Proceedings of the Royal Society of London Series B *276*, 13-19.

S40.   Nowak, M.A., Tarnita, C.E., and Wilson, E.O. (2010). The evolution of eusociality. Nature *466*, 1057-1062.

S41.   Doebeli, M. (2010). Inclusive fitness is just bookkeeping. Nature *467*, 661.

S42.   van Veelen, M. (2009). Group selection, kin selection, altruism and cooperation: when inclusive fitness is right and when it can be wrong. Journal of Theoretical Biology *259*, 589-600.

S43.   van Veelen, M., Garcia, J., Sabelis, M.W., and Egas, M. (2012). Group selection and inclusive fitness are not equivalent; the Price equation versus models and statistics. Journal of Theoretical Biology *299*, 64-80.

S44.   Traulsen, A. (2010). Mathematics of kin- and group-selection: formally equivalent? Evolution *64*, 316-323.

S45.   Damore, J.A., and Gore, J. (2012). Understanding microbial cooooperation. Journal of Theoretical Biology *299*, 31-41.

S46.   Dawkins, R. (1978). Replicator selection and the extended phenotype. Zeitschrift für Tierpsychologie *47*, 61-76.

S47.   Grafen, A. (2006). Optimisation of inclusive fitness. Journal of Theoretical Biology *238*, 541-563.

S48.   Grafen, A. (2007). The formal Darwinism project: a mid-term report. Journal of Evolutionary Biology *20*, 1243-1254.

S49.   Grafen, A. (1985). A geometric view of relatedness. Oxford Surv. Evol. Biol. *2*, 28-89.

S50.   Grafen, A. (1982). How not to measure inclusive fitness. Nature *298*, 425-426.

S51.   Gardner, A., and West, S.A. (2010). Greenbeards. Evolution *64*, 25-38.

S52.   Queller, D.C. (1984). Kin selection and frequency dependence: a game theoretic approach. Biological Journal of the Linnean Society *23*, 133-143.

S53. Queller, D.C. (1985). Kinship, reciprocity and synergism in the evolution of social behaviour. Nature *318*, 366-367.

S54. Queller, D.C. (2011). Expanded social fitness and Hamilton's rule for kin, kith, and kind. Proc. Natl. Acad. Sci. U. S. A. *108*, 10792-10799.

S55. Riley, M.A., and Wertz, J.E. (2002). Bacteriocins: evolution, ecology and application. A. Rev. Microbiol. *56*, 117-137.

S56. Biernaskie, J.M., West, S.A., and Gardner, A. (2011). Are greenbeards intragenomic outlaws? Evolution *65*, 2729-2742.

S57. McGinty, S.E., Rankin, D.J., and Brown, S.P. (2011). Horizontal gene transfer and the evolution of bacterial cooperation. Evolution *65*, 21-32.

S58. McGinty, S.E., and Rankin, D.J. (2012). The evolution of conflict resolution betwen plasmids and their bacterial hosts. Evolution *66*, 1662-1670.

S59. Nogueira, T., Rankin, D.J., Touchon, M., Taddei, F., Brown, S.P., and Rocha, E.P.C. (2009). Horizontal gene transfer of the secretome drives the evolution of bacterial cooperation and virulence. Current Biology *19*, 1683-1691.

S60. Platt, T.G., Bever, J.D., and Fuqua, C. (2012). A cooperative virulence plasmid imposes a high fitness cost under conditions that induce pathogenesis. Proceedings of the Royal Society of London Series B *279*, 1691-1699.

S61. Smith, J. (2001). The social evolution of bacterial pathogenesis. Proc. Roy. Soc. Lond. B *268*, 61-69.