

ORIGINAL ARTICLE

Genomic imprinting and the units of adaptation

A Gardner

Two guiding principles identify which biological entities are able to evolve adaptations. Williams' principle holds that, in order for an entity to evolve adaptations, there must be selection between such entities. Maynard Smith's principle holds that, in order for an entity to evolve adaptations, selection within such entities must be absent or negligible. However, although the kinship theory of genomic imprinting suggests that parent-of-origin-specific gene expression evolves as a consequence of natural selection acting between—rather than within—individuals, it evades adaptive interpretation at the individual level and is instead viewed as an outcome of an intragenomic conflict of interest between an individual's genes. Here, I formalize the idea that natural selection drives intragenomic conflicts of interest between genes originating from different parents. Specifically, I establish mathematical links between the dynamics of natural selection and the idea of the gene as an intentional, inclusive-fitness-maximizing agent, and I clarify the role that information about parent of origin plays in mediating conflicts of interest between genes residing in the same genome. These results highlight that the suppression of divisive information may be as important as the suppression of lower levels of selection in maintaining the integrity of units of adaptation.

Heredity (2014) **113**, 104–111; doi:10.1038/hdy.2013.128; published online 5 February 2014

Keywords: gene; group selection; kin selection; inclusive fitness; multi-level selection; organism

Is nayne in warld at scaithis ma do mar
Than weile trastyt in borne familiar

Hary, *The Wallace*

INTRODUCTION

Natural selection gives rise to biological adaptation: the appearance of design in the living world (Gardner, 2009). Usually, adaptive design is recognized at the level of the individual organism (Darwin, 1859), where it functions to maximize the individual's inclusive fitness (Hamilton, 1964). And, indeed, the individual-level adaptationist programme has been extraordinarily successful, for example, in the field of behavioural ecology (Davies *et al.*, 2012). However, adaptation may also manifest at other levels of biological organization, giving rise to ideas of 'selfish genes' (Hamilton, 1972; Dawkins, 1976, 1978; Burt and Trivers, 2006; Gardner and Welch, 2011) and 'superorganisms' (Wheeler, 1911; Wilson and Sober, 1989; Gardner and Grafen, 2009; Hölldobler and Wilson, 2009).

Two guiding principles identify which biological entities are able to evolve adaptations, and under which circumstances. Williams' principle holds that, in order for an entity to evolve adaptations, there must be selection *between* such entities (Williams, 1966; Sober and Wilson, 2011; Gardner, 2013). For example, group-level adaptation requires the operation of selection *between* groups. Maynard Smith's principle holds that, in order for an entity to evolve adaptations, selection *within* such entities must be absent or negligible (Maynard Smith, 1987, 1988; Gardner, 2013). For example, group-level adaptation requires that within-group selection be abolished. Formal justification for both of these principles is provided by Gardner and Grafen (2009).

However, these guiding principles do not appear to provide a complete account of the units of adaptation. According to the kinship theory of genomic imprinting (Haig, 2002), parent-of-origin-specific gene expression evolves as a consequence of natural selection acting between—rather than within—organisms. Yet, it evades adaptive interpretation at the level of the individual organism, and is instead viewed as an outcome of intragenomic conflict between the individual's maternally and paternally derived genes. The kinship theory has been empirically successful, correctly predicting the direction of imprint for genes that are of medical importance (Haig, 2002; Burt and Trivers, 2006). However, the idea that selection operating on such genes acts according to a design objective of fitness maximization at the gene level, and not at the individual level, remains to be formalized.

Here, I formalize the idea that natural selection may drive intragenomic conflicts of interest between genes originating from different parents. Specifically, I establish mathematical links between the dynamics of natural selection, captured by Price's (1970, 1972) equation, and the idea of the gene as an inclusive-fitness-maximizing agent, captured by an optimization program (Gardner and Welch, 2011), in the context of a population in which half of all genes are maternal in origin and half are paternal in origin. I then consider whether genes residing in the same individual but differing in their parent of origin may also differ in their maximands, yielding scope for intragenomic conflict of interest. In particular, I investigate the importance of a gene having information about its parent of origin in mediating this conflict of interest. My goal for this analysis is not to offer strikingly novel insights about genomic imprinting, but rather to formalize existing ideas and to illuminate more general principles of Darwinian adaptation.

GENES AS INCLUSIVE-FITNESS MAXIMIZERS

Overview

This section of the paper is concerned with justifying the view that genes maximize their inclusive fitness, and clarifying the meaning of this quantity. I build upon the analysis of Gardner and Welch (2011), who considered a simple model in which all genes are equivalent aside from the alleles carried by themselves and by those with whom they interact socially (that is, absence of class structure) that disallowed parent-of-origin specific phenotypes. Here, I extend that model to consider that there are two classes of gene: those deriving from the individual's mother and those deriving from the individual's father. First, I describe a population model of genes, their phenotypes and their fitnesses. Second, I use this population model to derive a statement about the action of natural selection, within the framework of Price's equation. Third, I use the population model to formalize statements about genes as inclusive-fitness maximizers, within the framework of the optimization program. Fourth, I derive mathematical connections between dynamics and optimization to justify the view that genes maximize their inclusive fitness. This sets the stage for a formal analysis of intragenomic conflict in the next section of the paper.

Population model

I consider a very large, but finite, population of diploid, sexual, hermaphrodite individuals, each assigned a unique index $i \in I$. For ease of exposition, I consider unicellular individuals (although this is not strictly necessary, as the mathematical analysis does allow for undifferentiated unicells to be grouped into multicellular assemblages). For simplicity, I assume that individuals have only one genetic locus and, within each individual, I assign each of the two genes a class $k \in K = \{m, p\}$ according to whether it is maternal or paternal in origin (each individual having one maternally derived gene and one paternally derived gene at the locus of interest). Accordingly, every gene in the population is uniquely identifiable by its carrier's index i and its own class k . For simplicity, I assume discrete generations. In addition, for ease of exposition, I assume that generations are non-overlapping (although the mathematical analysis does allow for overlapping generations, with a gene's survival to the next generation being exactly equivalent to it producing an extra replica). For simplicity, I assume the total number of genes is the same in every generation.

Every gene is assigned an allelic type $a \in A$, according to its nucleic acid sequence. Thus, the allele associated with the class- k gene in individual i is denoted a_{ik} . Every allele is associated with a genic value, given by $g = \mathcal{G}(a)$, where \mathcal{G} is the genotype function. Thus, the genic value associated with the class- k gene in individual i is denoted g_{ik} . The assignment of genic values to alleles is completely arbitrary. For example, $g = 1$ might be assigned to a particular allele and $g = 0$ to all the others, so that the population average genic value $E(g)$ will represent the frequency of the allele in the population. Alternatively, the genic value might reflect the average effect of the allele with respect to a phenotype of interest (Fisher, 1941).

I will consider two versions of the population model. In the first version, phenotypes may be mediated by class as well as allele, so that every allele is associated with two (potentially, though not necessarily, different) phenotypes, one for when the gene is maternal in origin and one for when the gene is paternal in origin. Thus, the class- k phenotype is given by $\pi_k = \mathcal{P}(a; k) \in P$, where \mathcal{P} is the phenotype function and P is the set of possible phenotypes. Note that I have assumed that both maternally and paternally derived genes may adopt any phenotype from the same set, P . In addition, I assume that there is

at least one allele that encodes any given pair of phenotypes (*cf.* Haig, 2000a). In the second version of the model, phenotypes are mediated only by allele and not by class. Thus, the class- k phenotype is given by $\pi_k = \mathcal{P}(a) \in P$. In both versions of the model, the phenotype associated with the class- k gene in individual i is denoted π_{ik} . All phenotypes are defined sufficiently proximately (for example, as a level of gene expression) that they may reasonably be considered under the sole control of a single gene.

Finally, a gene's fitness is given by the expected number of (mutated as well as non-mutated) replicas it leaves to the next generation. This expectation is taken over future uncertainty, which is not explicitly handled here (see Gardner and Welch, 2011 for details), and is jointly determined by a term $\tilde{\pi}$ that describes the phenotypic composition of the gene's social partners (including itself) and a term Π that describes the phenotypic composition of the entire population. The $\tilde{\pi}$ term may vary from gene to gene, as different genes may have different social partners, but the Π term is common to all genes in any given generation. This allows a formal separation of social versus global frequency-dependence effects (Gardner and Welch, 2011). Specifically, fitness is given by $w = \mathcal{W}(\tilde{\pi}; \Pi)$, where \mathcal{W} is the fitness function. Thus, the fitness of the class- k gene in individual i is denoted w_{ik} . I assume that, owing to fair meiosis, the expected fitness of a maternally derived gene is constrained to be equal to that of its paternally derived counterpart, that is, $w_{im} = w_{ip}$. This rules out the possibility of selection within individuals (for example, meiotic drive).

Dynamics of selection

Natural selection is defined as the population genetic change that is driven by systematic differences in individual fitness. This change may be captured using Price's (1970, 1972) equation:

$$\Delta_S E(g) = \text{cov}(w, g). \quad (1)$$

In words, the change in the average genic value that is ascribable to natural selection (left hand side) is equal to the covariance of fitness and genic value, taken over all genes in the population (right hand side). The idea that personal fitness (also termed 'neighbour-modulated fitness'; Hamilton, 1964) is a crucial covariate of genic value in the Price equation has led to it being termed a 'target of selection' (Grafen, 2002).

Importantly, I have not assumed that nonselective evolutionary change, owing to differences between parent and offspring genic values (that is, transmission effects), and/or deviations of realized fitness from its expected value (that is, random drift) are absent. Rather, I have simply set them aside as not of interest for the purpose of the present analysis. Briefly, this is because only selection—and not non-Darwinian factors such as mutation and random drift—is expected to give rise to a design rationale recognizable as biological adaptation. Gardner and Welch (2011) provide further details and discussion.

Maximization of inclusive fitness

I now turn from the dynamics of selection to the idea of the gene as a fitness-maximizing agent. This requires a stronger notion of fitness than simply 'target of selection'. Specifically, although the gene's personal fitness functions as an adequate covariate in the dynamical Equation (1), it does not follow that genes will act to maximize this quantity. This is because they do not have sole control of their personal fitness, insofar as this is also mediated by their social partners. Put another way, correlations between genic value and personal fitness may arise for reasons other than direct causation, and those genes that are favoured by natural selection are not necessarily

those that causally increase their personal fitness. What is needed is a reformulation of fitness that puts the focal gene in complete causal control. This is achieved by inclusive fitness (Hamilton, 1964; Gardner and Welch, 2011).

First, following the approach of Gardner and Welch (2011), I decompose a gene's personal fitness into baseline (B), additive (A) and non-additive (N) components:

$$\mathcal{W}(\tilde{\pi}; \Pi) = \mathcal{W}_B(\Pi) + \sum_J \sum_K \mathcal{W}_A(\pi_{jk}; j, k, \Pi) + \mathcal{W}_N(\tilde{\pi}; \Pi). \quad (2)$$

Here, the quantity $\mathcal{W}_A(\pi_{jk}; j, k, \Pi)$ is the additive contribution to a focal gene's personal fitness that is ascribed to the phenotype π_{jk} adopted by the class- k gene in its role- j social partner, in the context of a population with composition Π (Gardner and Welch, 2011). Note that the focal gene's own class does not appear in this calculation, as the personal fitness of a maternally derived gene is constrained to be the same as that of the paternally derived gene residing at the same locus in the same individual, and hence this quantity is independent of the gene's class.

Second, following the instructions of Hamilton (1964), I construct the gene's inclusive fitness by adding to baseline fitness the additive effects that the focal gene has upon the personal fitness of its social partners, weighting each increment or decrement by the appropriate coefficient of relatedness. This yields:

$$\mathcal{H}(\pi_l; l, \Pi) = \mathcal{W}_B(\Pi) + \sum_J \sum_K \mathcal{W}_A(\pi_l; j, l, \Pi) r_{kl}. \quad (3)$$

where r_{kl} is the genetic relatedness between a class- k recipient gene and the class- l gene acting in its role- j social partner, calculated from the perspective of the actor gene (Gardner and Welch, 2011).

Note that although I have defined inclusive fitness in terms of additive effects, I have not assumed that nonadditive fitness effects are absent from the model. Rather, they simply do not contribute to the computation of inclusive fitness. Later in this section, I will show that, despite this neglect of nonadditive effects, the inclusive fitness approach does provide a full account of the action of natural selection (see also Gardner and Welch, 2011). Also note that the maternally and paternally derived genes residing in the same individual are constrained to have the same fitness, and this is reflected in the absence of an argument k describing the recipient's class in the additive fitness effect \mathcal{W}_A .

Defining $r_{jl} = \frac{1}{2} r_{mjl} + \frac{1}{2} r_{pjl}$ as the relatedness between a recipient gene of unknown class and the class- l gene acting in its role- j social partner, Equation (3) may be rewritten as:

$$\mathcal{H}(\pi_l; l, \Pi) = \mathcal{W}_B(\Pi) + 2 \sum_J \mathcal{W}_A(\pi_l; j, l, \Pi) r_{jl}. \quad (4)$$

That inclusive fitness is under the sole control of the focal gene is made explicit in the fact that it takes as its argument only that gene's phenotype (plus the gene's class and the population context shared in common with all genes; Gardner and Welch, 2011) and not the whole list of phenotypes expressed by the gene's social partners. Accordingly, it is logically possible to regard the gene as striving to maximize its inclusive fitness (although so far, I have provided no further justification for doing so). Indeed, the idea that a class- l gene employs its phenotype $\pi_l \in P$ as a strategy in pursuit of the objective of inclusive-fitness maximization may be formally captured as an optimization program:

$$\pi_l \max_{\pi_l \in P} \mathcal{H}(\pi_l; l, \Pi). \quad (5)$$

The optimization program (5) not only provides a formal statement of the idea that a gene strives to maximize its inclusive

fitness, but also permits formal definitions of phenotypic optimality and suboptimality. Specifically, an optimal phenotype π_l^* is one that maximizes the gene's inclusive fitness within the constraints imposed by the strategy set, that is, $\pi_l^* \in P: \mathcal{H}(\pi_l^*; l, \Pi) \geq \mathcal{H}(\pi_l; l, \Pi) \forall \pi_l \in P$. And a suboptimal phenotype π_l° is one that fails to maximize inclusive fitness within the constraints imposed by the strategy set, that is, $\pi_l^\circ \in P: \exists \pi \in P: \mathcal{H}(\pi; l, \Pi) > \mathcal{H}(\pi_l^\circ; l, \Pi)$.

The inclusive fitness function (4) and corresponding optimization program (5) are appropriate if the focal gene knows its own class, and this is reflected in the presence of an argument l describing its class in the inclusive fitness function \mathcal{H} . However, if the gene is ignorant of its own class, then the appropriate inclusive fitness will be given by:

$$\mathcal{H}(\pi; \Pi) = \mathcal{W}_B(\Pi) + 2 \sum_J \mathcal{W}_A(\pi; j, \Pi) r_j, \quad (6)$$

where $r_j = \frac{1}{2} r_{jm} + \frac{1}{2} r_{jp}$ is the relatedness between a recipient gene of unknown class and an actor gene of unknown class in its role- j social partner. The corresponding idea that the gene employs its phenotype $\pi \in P$ as a strategy in pursuit of the objective of inclusive-fitness maximization may also be formally captured as an optimization program:

$$\pi \max_{\pi \in P} \mathcal{H}(\pi; \Pi) \quad (7)$$

Again, the optimization program (7) permits formal definitions of phenotypic optimality and suboptimality: an optimal phenotype π^* is one that maximizes inclusive fitness within the constraints imposed by the strategy set, that is, $\pi^* \in P: \mathcal{H}(\pi^*; \Pi) \geq \mathcal{H}(\pi; \Pi) \forall \pi \in P$; and a suboptimal phenotype π° is one that fails to maximize inclusive fitness within the constraints imposed by the strategy set, that is, $\pi^\circ \in P: \exists \pi \in P: \mathcal{H}(\pi; \Pi) > \mathcal{H}(\pi^\circ; \Pi)$. More generally, the distinction between optimization programs (5) and (7) formally captures the idea that a gene's intention may depend on the information that it has at its disposal (in this case, concerning its parent of origin).

Inclusive fitness as target of selection and maximand

I derived an expression (1) for the dynamics of natural selection, in which personal fitness is given as a target of selection. I then showed that inclusive fitness, rather than personal fitness, is logically of the correct form to be considered a maximand of an intentional gene. I now show that inclusive fitness also functions as a target of selection, and I derive mathematical connections between the dynamics of selection and the maximization of inclusive fitness to provide formal justification for the idea that natural selection acts according to a design objective of inclusive-fitness maximization at the gene level.

Considering the version of the population model in which a gene's phenotype is determined both by its allele and its class, I am able to derive a link between the dynamical Equation (1) and the optimization program (5), such that the target of selection is the inclusive-fitness maximand appropriate for when the gene is privy to information about its parent of origin:

$$\Delta_S E(g) = \frac{1}{2} \sum_L \text{cov}(\mathcal{H}(\pi_l; l, \Pi), g_l), \quad (8)$$

(see Appendix for derivation). More generally, I derive six mathematical correspondences between the dynamical Equation (1) and the optimization program (5), equivalent to those derived by Gardner and Welch (2011) for a nonclass-structured population (Table 1; derivations in Appendix). These correspondences provide formal justification for the idea that natural selection acts according to a design objective of inclusive-fitness maximization at the gene level, and specifically for the idea that a gene's intentions may depend on its

Table 1 Correspondences between dynamical and purposeful accounts of gene-level adaptation

Numeral	Correspondence
I	If all agents are optimal, there is no scope for selection (no expected change in the average of any genic value)
II	If all agents are optimal, there is no potential for positive selection (no introduced allele can increase from rarity due to selection)
III	If all agents are suboptimal, but equally so, there is no scope for selection
IV	If all agents are suboptimal, but equally so, there is potential for positive selection (there exists an allele that, if introduced, can increase from rarity due to selection)
V	If agents vary in their optimality, there is scope for selection, and change in the average of all genic values is given by their average within-class covariance with attained maximand
VI	If there is neither scope for selection nor potential for positive selection, all agents are optimal

Formal justification for the 'gene as fitness-maximizing agent' view rests upon the ability to translate this way of thinking into formal population genetics, and vice versa. Derivations are provided in Appendix.

parent of origin when it has the option of employing different phenotypes according to whether it is of maternal or paternal origin.

Considering the version of the population model in which a gene's phenotype is determined solely by its allele and is not mediated by its class, I am able to derive a link between the dynamical Equation (1) and the optimization program (7), such that the target of selection is the inclusive-fitness maximand appropriate for when the gene is ignorant of its parent of origin:

$$\Delta_S E(g) = \text{cov}(\mathcal{H}(\pi; \Pi), g), \quad (9)$$

(see Appendix for derivation). More generally, I derive six mathematical correspondences between the dynamical Equation (1) and the optimization program (5), equivalent to those described above and by Gardner and Welch (2011) for a nonclass-structured population (Table 1; derivations in Appendix). These correspondences provide formal justification for the idea that natural selection acts according to a design objective of inclusive-fitness maximization at the gene level, and specifically for the idea that a gene's intentions do not depend on its parent of origin when it does not have the option of employing different phenotypes according to whether it is of maternal or paternal origin.

INTRAGENOMIC CONFLICT

In the previous section, I provided formal justification for the idea that genes maximize their inclusive fitness in the context of a population structured into genes of maternal origin and genes of paternal origin. Moreover, I showed that if a gene's phenotype may be mediated by its parent of origin and not only by its allele, then the gene's inclusive-fitness agenda will be such that it appears to have information about its parent of origin. Having formalized the intentionality of genes, I now consider how their agendas give rise to intragenomic conflicts of interest.

First, I consider the scenario where a gene's phenotype is not mediated by its parent of origin, but determined solely by its allele. Accordingly, genes optimize as if they are ignorant as to their parent of origin, that is, according to optimization program (7). That is, they strive to maximize the inclusive fitness function (6). How does this inclusive fitness function differ for different genes? If the two genes are in two different individuals, then they potentially interact with different social partners when acting within any particular role, j . Importantly, they may share individual social partners (including each other), but as they interact with them in potentially different roles (different j 's), they are liable to value these same social partners with different coefficients of relatedness (different r_j 's)—and this gives potential for a conflict of interests. However, if the two genes are in the same individual—specifically, the maternally and the paternally derived gene at the locus of interest—then they interact with exactly

the same set of social partners, and always in the same roles (same j), and hence they will always value these social partners with the same coefficients of relatedness (same r_j 's)—and this yields no potential for a conflict of interests.

Second, I consider the scenario where a gene's phenotype is mediated by its parent of origin as well as its allele. Accordingly, genes optimize as if they are privy to information about their parent of origin, that is, according to optimization program (5). That is, they strive to maximize the inclusive fitness function (4). How does this inclusive fitness function differ for different genes? If the two genes are in two different individuals, then—once again—they potentially interact with the same social partners but in the context of different roles, j , and this gives potential for a conflict of interests. However, if the two genes are in the same individual—specifically, the maternally and the paternally derived genes at the locus of interest—then, although they interact with exactly the same set of social partners and always in the same roles (same j), they need not always value these same social partners with the same coefficients of relatedness, because relatedness also depends on the gene's class, and this is different for genes of maternal versus paternal origin (different r_{ji} 's, that is, r_{jm} versus r_{jp}). Accordingly, if individuals may be differentially related to their social partners via their two parents, then their maternally derived genes may disagree with their paternally derived genes about the valuation of those social partners, leading to intragenomic conflicts of interest.

The formal separation of selective dynamics from inclusive-fitness agenda also helps to clarify the distinction between the conflict of interest between an individual's maternally paternally derived genes versus what has been termed 'parental antagonism' (Haig, 1997). A parentally antagonistic allele is one that enjoys a selective advantage when inherited via one parent and suffers a selective disadvantage when inherited via the other parent; a scenario that is captured in the dynamical Equation (8) whenever the covariance of relative fitness and the genic value associated with this allele is positive among genes originating from one parent, and negative among genes originating from the other parent. The resulting interference in the overall genetic change is properly viewed as a constraint to adaptation rather than a true intragenomic conflict. The very existence of alleles with parent-of-origin-specific expression means that it is possible for there to be intragenomic conflict between maternally and paternally derived genes (that is, different inclusive fitness agendas) without there being parental antagonism (that is, selection covariance of each class having opposite sign). In addition, one may imagine scenarios in which there would be parental antagonism without intragenomic conflict; for example, if there were a cooperative division of labour between the maternally and paternally derived genomes,

with certain alleles being good at fulfilling the adaptive role of one class but not the other.

DISCUSSION

Application of Williams' and Maynard Smith's principles of adaptation suggests that any biological entity may evolve adaptations so long as there is selection between, and not within, such entities. However, parent-of-origin-specific gene expression poses a problem for this simplistic view of adaptation. The kinship of genomic imprinting suggests that such a pattern of gene expression is not optimized according to the interests of the individual, despite being driven by selection acting between rather than within individuals, and must instead be interpreted as a gene-level adaptation. Here, I have provided formal support for the suggestion that natural selection acting between individuals may drive intragenomic conflicts of interest between genes that differ in their parent of origin. I have explicitly connected an optimization program that formally captures the idea of the gene as an inclusive-fitness maximizer with Price's equation of selection dynamics, and shown that information about parent of origin may lead different genes in the same individual to have different social agendas—owing to their differential valuation of maternal versus paternal relations—despite the absence of selection within the individual. This understanding about the source of intragenomic conflicts between genes with different parent of origin is, of course, not new (Haig, 2000a provides an excellent overview). However, the present analysis is the first to explicitly examine the divergent agendas of maternally versus paternally derived genes that formally emerge from the dynamics of natural selection.

Limitations of the model

For the purpose of simplicity and clarity, my analysis has been deliberately limited in a number of respects. First, my focus has been on uncovering the existence of conflicts of interest between genes and not the details of the resolution of this conflict. Accordingly, I have investigated how genes with different parents of origin may prefer to adopt divergent phenotypes in pursuit of the maximization of their inclusive fitness, which is a prerequisite for genomic imprinting, without considering the interplay of potentially multiple genes—including imprinters as well as imprinted (Burt and Trivers, 1998; Haig, 2000a)—that is required to actually bring about parent-of-origin differences in gene expression levels. Second, in order to avoid class structure beyond the designation of genes as maternal versus paternal in origin, I have assumed a population of hermaphrodites. Separation of sexes would require four classes of genes, designated according to their carrier's sex as well as parent of origin (*cf.* Haig, 1997). This simplification rules out some of the demographic factors that have been suggested to drive the evolution of genomic imprinting. For example, sex-biased dispersal provides a means of making an individual more related to her neighbours through her parent of the less-dispersing sex (Haig, 2000b; Úbeda and Gardner, 2010, 2011, 2012; Úbeda *et al.*, 2014). However, a difference in the dispersal rates of female versus male gametes in hermaphrodite broadcast spawners could yield a similar result.

Tokens versus types

I have followed the approach of Gardner and Welch (2011) in identifying the genic agent as a scrap of nucleic acid (that is, a physical object or 'token') rather than as an allelic variant (that is, an abstract category or 'type'). Genes are often described as being 'selfish' (Dawkins, 1976), but genic tokens, as maximizers of inclusive fitness, may equally be altruistic or spiteful. There is a sense in which allelic

types may be informally considered as 'selfish', in that the alleles that are evolutionarily successful are (trivially) those that increase their representation in the population. However, this is true of alleles that predominate by mutation pressure or random drift, and not just by natural selection, so this idea of the selfish allele has a distinctly non-Darwinian flavour. Moreover, it is tokens rather than types that directly manifest adaptation, as they are real-world objects rather than residents of a Platonic realm of abstractions. Furthermore, the idea that an "offspring's paternally derived genes have different interests from the offspring's maternally derived genes" (Haig, 2002) can only make sense if genes are tokens rather than types, because it is the genic tokens rather than the allelic types that have parents of origin. In any case, if the allele-as-maximizing-agent analogy can be mathematically formalized, this remains to be done.

Haig (2012) suggests that this genic token approach may only be valid owing to the artificial assumption of no class structure in the model of Gardner and Welch (2011) that "collapses everything a gene might do into a single token that can be considered the recipient of its own effects" (Haig, 2012). This is incorrect, and here I have shown that the approach extends to a model in which genes are separated into nonequivalent maternal-origin versus paternal-origin classes. As an alternative, Haig (2012) suggests that the 'strategic gene' is better viewed as a collective of identical-by-descent tokens. This has some validity if we consider a collective that is guaranteed to be identical by descent, according to the information that is available to every member of the collective; for example, all the maternally derived genes residing at a particular locus within the body of an individual organism. Indeed, the formal justification for the view that a genetically uniform collective may be viewed as an adaptive agent is provided by Gardner and Grafen (2009).

However, Haig (2012) intends the strategic gene to encompass all the identical-by-descent copies in all the focal individual's social partners (which collapses everything a gene might do into a single token that can be considered the recipient of its own effects). This is problematic, because a focal genic token does not necessarily *know* which other genic tokens are identical by descent. Rather, from the perspective of any given gene, there will generally be a wide collection of genic social partners to which it has a nonzero probability of identity by descent. Consequently, the genes in this diffuse cloud will not be selectively favoured to consider each other as 'self', but rather to value each other according to their probability of identity by descent—just as individual organisms who engage in social interaction with nonidentical relatives do not consider them as 'self' but rather value them according to probabilities of sharing alleles in common (Hamilton, 1964; Grafen, 2006). Moreover, Haig's notion of the strategic gene is also incompatible with describing a gene as being maternal or paternal in origin as, within any collective of identical-by-descent tokens stretching across multiple individuals, different tokens are liable to have different parents of origin.

Principles of adaptation

More generally, the present analysis clarifies that, as a general principle of adaptation, the suppression of divisive information may be just as important as the suppression of lower levels of selection (Maynard Smith's principle) in sustaining the integrity of a unit of adaptation (Bourke, 2011; Gardner, 2013). This principle was overlooked in the analysis of Gardner and Grafen (2009) that examined conditions under which social groups may be considered adaptive units in their own right, owing to their crucial assumption that all social interaction takes place within and not between social groups. Thus, in the 'repression of competition' model of Gardner and Grafen

(2009), in which group mates are genetically heterogeneous but constrained to have the same expected fitness (for example, genes in a genome under fair meiosis; Leigh, 1977), the only way for the lower-level units to increase their inclusive fitness is to increase the fitness of their social group. However, if social interactions also extend between groups, then inclusive fitness may be increased at the expense of one's group if it provides benefits to one's relatives in other groups—who might not be relatives of one's group mates. This highlights that altruistic behaviour, although often conceived as a step along the road to a major transition in individuality, may actually reignite conflicts of interest between lower-level units that erode the integrity of existing units of adaptation (Gardner, 2013).

These insights also clarify the error in the application of William's principle by Sober and Wilson (2011) that led them to conclude that individuals in clonal groups are not units of adaptation on the basis that there is no selection at the individual level in such scenarios. The same reasoning would suggest that imprinted genes are not adaptive agents, and that the adaptive rationale for genomic imprinting must be sought at the individual level, because there is no selection at the gene level—a suggestion that the present analysis has formally refuted. Sober and Wilson's (2011) error is to confuse 'individual selection' with 'within-group selection' (and, accordingly, 'gene selection' with 'within-individual selection'), so that the absence of within-group selection in the context of clonal groups appears to preclude adaptation at the individual level. However, the formal analysis of Gardner and Grafen (2009) clarifies that the proper notion of 'individual selection' is selection among all the individuals in the population, and not just those within a particular group. Consequently, there is selection between individuals even in populations with clonal groups, because an individual in one group may be genetically distinct (and fitter) than an individual in a different group (Gardner, 2014). Thus, it is possible to simultaneously view the group and the individual as units of adaptation, just as a honeybee retains her adaptive, inclusive-fitness-maximizing integrity even though her colony may qualify as a superorganism in its own right. Moreover, this clarifies the importance of mathematical (rather than just verbal) theories of adaptation (contra Sober and Wilson, 2011), as words are slippery and may confuse more than they clarify.

CONFLICT OF INTEREST

The author declares no conflict of interest.

ACKNOWLEDGEMENTS

I thank David Shuker, John Welch and three anonymous referees for helpful comments and discussion, and the Royal Society for funding.

APPENDIX

Phenotype mediated by allele and class

Noting that $g = \frac{1}{2} \sum_L g_L$, the action of natural selection can be written as:

$$\Delta_S E(g) = \frac{1}{2} \sum_L \text{cov}(w, g_L). \quad (\text{A1})$$

Substituting Equation (2) into Equation (A1) obtains:

$$\Delta_S E(g) = \frac{1}{2} \sum_J \sum_K \sum_L \text{cov}(\mathcal{W}_A(\pi_{jk}; j, k, \Pi), g_L). \quad (\text{A2})$$

Using the fact that $\text{cov}(x, y) = \beta(y, x) \text{cov}(x, x)$, where $\beta(y, x)$ is the least-squares linear regression of y against x , Equation (A2) can be

- Bourke AFG (2011). *Principles of Social Evolution*. Oxford University Press.
- Burt A, Trivers R (1998). Genetic conflicts in genomic imprinting. *Proc R Soc Lond B* **265**: 2393–2397.
- Burt A, Trivers R (2006). *Genes in Conflict*. Harvard University Press.
- Darwin CR (1859). *The Origin of Species*. John Murray: London.
- Davies NB, Krebs JR, West SA (2012). *An Introduction to Behavioural Ecology*, 4th edn Wiley-Blackwell.
- Dawkins R (1976). *The Selfish Gene*. Oxford University Press.
- Dawkins R (1978). Replicator selection and the extended phenotype. *Z Tierpsychol* **47**: 61–76.
- Fisher RA (1941). Average excess and average effect of a gene substitution. *Ann Eugen* **11**: 53–63.
- Gardner A (2009). Adaptation as organism design. *Biol Letts* **5**: 861–864.
- Gardner A (2013). Adaptation of individuals and groups. In: Bouchard F, Huneman P (eds) *From Groups to Individuals*. MIT Press: London, pp 99–116.
- Gardner A (2014). Life, the universe and everything. *Biol Philos* doi:10.1007/s10539-013-9417-8.
- Gardner A, Grafen A (2009). Capturing the superorganism: a formal theory of group adaptation. *J Evol Biol* **22**: 659–671.
- Gardner A, Welch JJ (2011). A formal theory of the selfish gene. *J Evol Biol* **24**: 1801.
- Grafen A (2002). A first formal link between the Price equation and an optimization program. *J Theor Biol* **217**: 75–91.
- Grafen A (2006). Optimization of inclusive fitness. *J Theor Biol* **238**: 541–563.
- Haig D (1997). Parental antagonism, relatedness asymmetries, and genomic imprinting. *Proc R Soc Lond B* **264**: 1657–1662.
- Haig D (2000a). The kinship theory of genomic imprinting. *Annu Rev Ecol Syst* **31**: 9–32.
- Haig D (2000b). Genomic imprinting, sex-biased dispersal, and social behaviour. *Ann NY Acad Sci* **907**: 149–163.
- Haig D (2002). *Genomic imprinting and kinship*. Rutgers University Press.
- Haig D (2012). The strategic gene. *Biol Philos* **27**: 461–479.
- Hamilton WD (1964). The genetical evolution of social behaviour. *J Theor Biol* **7**: 1–52.
- Hamilton WD (1972). Altruism and related phenomena, mainly in social insects. *Annu Rev Ecol Syst* **3**: 193–232.
- Hölldobler B, Wilson EO (2009). *The Superorganism: The Beauty, Elegance and Strangeness of Insect Societies*. WW Norton: London.
- Leigh EG Jr (1977). How does selection reconcile individual advantage? *Proc Natl Acad Sci USA* **74**: 4542–4546.
- Maynard Smith J (1987). Reply to Sober. In: Dupré J (ed) *The Latest on the Best: Essays on Evolution and Optimality*. MIT Press: London, pp 147–149.
- Maynard Smith J (1988). Evolutionary progress and the levels of selection. In: Nitecki MH (ed) *Evolutionary Progress*. University of Chicago Press: Chicago, pp 219–230.
- Price G R (1970). Selection and covariance. *Nature* **227**: 520–521.
- Price GR (1972). Extension of covariance selection mathematics. *Ann Hum Genet* **35**: 485–490.
- Sober E, Wilson DS (2011). Adaptation and natural selection revisited. *J Evol Biol* **24**: 462–468.
- Úbeda F, Gardner A (2010). A model for genomic imprinting in the social brain: juveniles. *Evolution* **64**: 2587–2600.
- Úbeda F, Gardner A (2011). A model for genomic imprinting in the social brain: adults. *Evolution* **65**: 462–475.
- Úbeda F, Gardner A (2012). A model for genomic imprinting in the social brain: elders. *Evolution* **66**: 1567–1581.
- Úbeda F, Ohtsuki H, Gardner A (2014). Intragenomic conflict over menopause. *Ecol Lett* **17**: 165–174.
- Wheeler WM (1911). The ant-colony as an organism. *J Morphol* **22**: 307–325.
- Williams GC (1966). *Adaptation and Natural Selection*. Princeton University Press.
- Wilson DS, Sober S (1989). Reviving the superorganism. *J Theor Biol* **136**: 337–356.

rewritten as:

$$\Delta_S E(g) = \frac{1}{2} \sum_J \sum_K \sum_L \beta(\mathcal{W}_A(\pi_{jk}; j, k, \Pi), g_{jk}) \text{cov}(g_{jk}, g_L). \quad (\text{A3})$$

Noting that $K = L = \{m, p\}$, Equation (A3) can be rewritten as:

$$\Delta_S E(g) = \frac{1}{2} \sum_J \sum_K \sum_L \beta(\mathcal{W}_A(\pi_{jl}; j, l, \Pi), g_{jl}) \text{cov}(g_{jl}, g_k), \quad (\text{A4})$$

or:

$$\Delta_S E(g) = \frac{1}{2} \sum_J \sum_K \sum_L \beta(\mathcal{W}_A(\pi_l; j, l, \Pi), g_l) r_{kjl} \text{cov}(g_l, g_l), \quad (\text{A5})$$

where $r_{kjl} = \text{cov}(g_k, g_{jl}) / \text{cov}(g_l, g_l)$. Finally, some rearrangement gives:

$$\Delta_S E(g) = \frac{1}{2} \sum_L \text{cov} \left(\sum_J \sum_K \mathcal{W}_A(\pi_i; j, l, \Pi) r_{kjl}, g_l \right), \quad (\text{A6})$$

which is equivalent to Equation (8) of the main text. This result allows the derivation of the following correspondences:

I. If all agents are optimal, there is no scope for selection

If all agents behave optimally according to optimization program (5), then $\mathcal{H}(\pi_i; l, \Pi) = \mathcal{H}(\pi_i^*; l, \Pi) \forall i \in I$ and $\forall l \in L$. Hence, from Equation (8), the action of selection is $\Delta_S E(g) = (1/2) \sum_L \text{cov}(\mathcal{H}(\pi_i; l, \Pi), g_l) = (1/2) \sum_L \text{cov}(\mathcal{H}(\pi_i^*; l, \Pi), g_l) = 0$.

II. If all agents are optimal, there is no potential for positive selection

Choose a population composition in the vicinity of Π such that a new allele $a' \in A$ is present at vanishingly low frequency and with negligible impact upon the genetic relatedness of social partners. Carriers of the allele a' have phenotype $\pi_i' = \mathcal{P}(a'; l)$, according to their class; and all other genes, being optimal according to optimization program (5), have phenotype π_m^* or π_p^* , according to their class. Assigning carriers of allele a' a genic value $g=1$ and carriers of all other alleles a genic value $g=0$ then, from Equation (8), the action of selection is $\Delta_S E(g) = (1/2) \sum_L \text{cov}(\mathcal{H}(\pi_i; l, \Pi), g_l) = (1/2) \sum_L E(g_l) (1 - E(g_l)) (\mathcal{H}(\pi_i'; l, \Pi) - \mathcal{H}(\pi_i^*; l, \Pi))$. As $\mathcal{H}(\pi_i'; l, \Pi) \leq \mathcal{H}(\pi_i^*; l, \Pi)$ for all $\pi' \in P$, then $\Delta_S E(g) \leq 0$.

III. If all agents are suboptimal, but equally so, there is no scope for selection

If all agents behave suboptimally but equally so according to optimization program (5), then $\mathcal{H}(\pi_i; l, \Pi) = \mathcal{H}(\pi_i^\circ; l, \Pi) \forall i \in I$ and $\forall l \in L$. Hence, from Equation (8), the action of selection is $\Delta_S E(g) = (1/2) \sum_L \text{cov}(\mathcal{H}(\pi_i^\circ; l, \Pi), g_l) = (1/2) \sum_L \text{cov}(\mathcal{H}(\pi_i; l, \Pi), g_l) = 0$.

IV. If all agents are suboptimal, but equally so, there is potential for positive selection

Choose a population composition in the vicinity of Π such that the allele a^* corresponding to the optimal pair of phenotypes π_m^* and π_p^* , according to optimization program (5), is present at vanishingly low frequency and with negligible impact upon the genetic relatedness of social partners. All other genes, being suboptimal according to optimization program (5), have phenotypes π_m° or π_p° , according to their class. Assigning carriers of the allele a^* a genic value $g=1$ and carriers of the other alleles a genic value $g=0$ then, from Equation (8), the action of selection is $\Delta_S E(g) = (1/2) \sum_L \text{cov}(\mathcal{H}(\pi_i; l, \Pi), g_l) = (1/2) \sum_L E(g_l) (1 - E(g_l)) (\mathcal{H}(\pi_i^*; l, \Pi) - \mathcal{H}(\pi_i^\circ; l, \Pi))$. As $\mathcal{H}(\pi_i^*; l, \Pi) > \mathcal{H}(\pi_i^\circ; l, \Pi)$, then $\Delta_S E(g) > 0$.

V. If agents vary in their optimality, there is scope for selection, and the change in the average of all genic values is given by their average within-class covariance with attained maximand

From Equation (8), $\Delta_S E(g) = (1/2) \sum_L \text{cov}(\mathcal{H}(\pi_i; l, \Pi), g_l)$.

VI. If there is neither scope for selection nor potential for positive selection, all agents are optimal

If all agents are optimal, there is neither scope for selection (I) nor potential for positive selection (II); if all agents are suboptimal but equally so, then there is no scope for selection (III) but there is potential for positive selection (IV); and if agents vary in their optimality, there is scope for selection (V). These exhaust the

possibilities in the optimization view; hence, if there is neither scope for selection nor potential for positive selection, all agents must be optimal.

Phenotype mediated by allele but not class

Substituting Equation (2) into Equation (1) obtains:

$$\Delta_S E(g) = \sum_J \sum_K \text{cov}(\mathcal{W}_A(\pi_{jk}; j, k, \Pi), g) \quad (\text{A7})$$

which, as $\mathcal{W}_A(\pi_{jk}; j, k, \Pi) = \mathcal{W}_A(\pi_j; j, \Pi)$ for all $k \in K$, is equivalent to:

$$\Delta_S E(g) = 2 \sum_J \text{cov}(\mathcal{W}_A(\pi_j; j, \Pi), g). \quad (\text{A8})$$

This may be written as:

$$\Delta_S E(g) = 2 \sum_J \beta(\mathcal{W}_A(\pi_j; j, \Pi), g) \text{cov}(g, g). \quad (\text{A9})$$

Note that $\beta(\mathcal{W}_A(\pi_j; j, \Pi), g) = \beta(\mathcal{W}_A(\pi_j; j, \Pi), g_j) \times \beta(g_j, g)$, $\beta(\mathcal{W}_A(\pi_j; j, \Pi), g_j) = \beta(\mathcal{W}_A(\pi_j; j, \Pi), g)$ and $\beta(g_j, g) = r_j$. Making these substitutions yields:

$$\Delta_S E(g) = 2 \sum_J \beta(\mathcal{W}_A(\pi_j; j, \Pi), g) r_j \text{cov}(g, g), \quad (\text{A10})$$

which may be written as:

$$\Delta_S E(g) = \text{cov} \left(2 \sum_J \mathcal{W}_A(\pi_j; j, \Pi) r_j, g \right), \quad (\text{A11})$$

which is equal to Equation (9) of the main text. This result allows the derivation of the following correspondences:

I. If all agents are optimal, there is no scope for selection

If all agents behave optimally according to optimization program (7), then $\mathcal{H}(\pi; \Pi) = \mathcal{H}(\pi^*; \Pi) \forall i \in I$. Hence, from Equation (9), the action of selection is $\Delta_S E(g) = \text{cov}(\mathcal{H}(\pi; \Pi), g) = \text{cov}(\mathcal{H}(\pi^*; \Pi), g) = 0$.

II. If all agents are optimal, there is no potential for positive selection

Choose a population composition in the vicinity of Π such that a new allele $a' \in A$ is present at vanishingly low frequency and with negligible impact upon the genetic relatedness of social partners. Carriers of the allele a' have phenotype $\pi' = \mathcal{P}(a')$ and all other genes, being optimal according to optimization program (7), have phenotype π^* . Assigning carriers of allele a' a genic value $g=1$ and carriers of all other alleles a genic value $g=0$ then, from Equation (9), the action of selection is $\Delta_S E(g) = \text{cov}(\mathcal{H}(\pi; \Pi), g) = E(g) (1 - E(g)) (\mathcal{H}(\pi'; \Pi) - \mathcal{H}(\pi^*; \Pi))$. As $\mathcal{H}(\pi'; \Pi) \leq \mathcal{H}(\pi^*; \Pi)$ for all $\pi' \in P$, then $\Delta_S E(g) \leq 0$.

III. If all agents are suboptimal, but equally so, there is no scope for selection

If all agents behave suboptimally but equally so according to optimization program (7), then $\mathcal{H}(\pi; \Pi) = \mathcal{H}(\pi^\circ; \Pi) \forall i \in I$. Hence, from Equation (9), the action of selection is $\Delta_S E(g) = \text{cov}(\mathcal{H}(\pi^\circ; \Pi), g) = \text{cov}(\mathcal{H}(\pi; \Pi), g) = 0$.

IV. If all agents are suboptimal, but equally so, there is potential for positive selection

Choose a population composition in the vicinity of Π such that the allele a^* corresponding to the optimal phenotypes π^* , according to optimization program (7), is present at vanishingly low frequency and with negligible impact upon the genetic relatedness of social partners. All other genes, being suboptimal according to optimization

program (7), have phenotype π° . Assigning carriers of the allele a^* a genic value $g=1$ and carriers of the other alleles a genic value $g=0$ then, from Equation (9), the action of selection is $\Delta_S E(g) = \text{cov}(\mathcal{H}(\pi; \Pi), g) = E(g)(1 - E(g)) (\mathcal{H}(\pi^*; \Pi) - \mathcal{H}(\pi^\circ; \Pi))$. As $\mathcal{H}(\pi^*; \Pi) > \mathcal{H}(\pi^\circ; \Pi)$, then $\Delta_S E(g) > 0$.

V. If agents vary in their optimality, there is scope for selection, and the change in the average of all genic values is given by their average within-class covariance with attained maximand

From Equation (9), $\Delta_S E(g) = \text{cov}(\mathcal{H}(\pi; \Pi), g)$. Note that as there is effectively only one class of genes here—that is, genes that are unaware of their parent of origin—then the average

within-class covariance is equal to the total covariance within that single class.

VI. If there is neither scope for selection nor potential for positive selection, all agents are optimal

If all agents are optimal, there is neither scope for selection (I) nor potential for positive selection (II); if all agents are suboptimal but equally so, then there is no scope for selection (III) but there is potential for positive selection (IV); and if agents vary in their optimality, there is scope for selection (V). These exhaust the possibilities in the optimization view; hence, if there is neither scope for selection nor potential for positive selection, all agents must be optimal.