Biol Philos (2014) 29:155–174
DOI 10.1007/s10539-013-9414-y

The formal darwinism project in outline

Alan Grafen

Accepted: 23 December 2013 / Published online: 25 January 2014

Abstract The formal darwinism project aims to provide a mathematical framework within which important fundamental ideas in large parts of biology can be articulated, including Darwin's central argument in *The Origin* (that mechanical processes of inheritance and reproduction can give rise to the appearance of design), modern extensions of evolutionary theory including ESS theory and inclusive fitness, and Dawkins' synthesis of them into a single structure. A new kind of argument is required to link equations of motion on the one hand to optimisation programs on the other, and a major point is that the biologist's concept of fitness maximisation is not represented by concepts from dynamical systems such as Lyapunov functions and gradient functions. The progress of the project so far is reviewed, though with only a brief glance at the rather complicated mathematics itself, and the centrality of fitness maximisation ideas to many areas of biology is emphasised.

Keywords Formal darwinism · Fitness maximisation · Behavioural ecology · Biological design · Population genetics · Natural selection

1 Introduction

The goal of the formal darwinism project is to construct a mathematical bridge between two of the many ways of studying natural selection. One approach is population genetics, in which models are constructed that trace the change over time of the frequencies of some defined set of genotypes. As biologists agree that changing genotype frequencies is how evolution happens, this approach is very firmly grounded so far as the process of natural selection is concerned. The other

A. Grafen (✉)
St John's College, Oxford, OX1 3JP, UK
e-mail: alan.grafen@sjc.ox.ac.uk

approach is based on the expectation of finding good design in nature: this stretches back at least to natural theology in the eighteenth century, and was invigorated and reinforced as a scientific approach by Darwin (1859) and later Fisher (1930). Design is about the outcome of natural selection, but suffers in comparison by being often less precise and less easily formulated in objective terms—the call of Williams (1966) for a new science of *teleonomy* has gone noticeably unanswered.

This paper aims to describe the formal darwinism project, but with a minimum of mathematics. As the project is highly technical, this does present difficulties, but I have tried not to abuse my double role as author of the mathematical arguments and their translator into words. A mathematical account would be very long, and unsuitable for this journal. The core papers are Grafen (1999, 2000, 2002, 2006a, 2006b) with exposition in Grafen (2007b, 2008).

Section 2 explains that the need for some bridging between the approaches has been recognised from some time, discusses previous attempts, delineates some of the difficulties, and sets out a prospectus of how bridging might look. Section 3 introduces the formal darwinism project with a minimum of mathematics, illustrates the slipperiness of one well-known ingredient—the so-called "Price Equation"—, states generic forms of the project's typical 'four links' between the two approaches, and ends with a sketch of the results of the project published by the end of 2012. Section 4 looks at why those four links are significant, even though they fall short of proving that fitness is always maximised. Section 5 sets the project briefly in context by discussing neighbouring literatures, which either assume or are frequently cited for demonstrating that fitness is maximised or, alternatively, purport to show that it is not. The discussion in Sect. 6 offers a sneak preview of significant work currently in the pipeline, and then tries to guess why this paper is being published in this journal.

2 Background

The design approach to studying evolution has a long and distinguished history. Borrowing from natural theology, but with a crucial twist, Darwin (1859) employed evidence of design in the natural world to argue for the existence of natural selection, and to explore its nature. Anatomy and morphology that linked form to function also took the design approach. Today, as molecular biologists choose to call some of their discoveries 'mechanisms', and ascribe 'functions' to enzymes, they use purposive language and so they also adopt the design approach. It is arguably impossible to undertake work in many areas of biology without doing so: purpose in explanations has great power, and attempts to do without it in ethology (for example, Kennedy 1992, reviews his earlier campaign in ethology as well as bringing in further subjects), have long ago been abandoned as unworkable. In parts of the subject susceptible to precision, the design approach has been formalised into a structure that includes a set of possible phenotypes, an evaluation of each phenotype in a number that represents how successful it is, and a presumption that we should see in nature the feasible phenotype with the highest evaluation. Following Herbert Spencer's possibly unfortunate coinage of 'the

survival of the fittest', the evaluation is often called fitness, and we will refer to the paradigm as fitness-maximisation.

Many areas of modern biology are founded on the design approach, including Behavioural Ecology (Davies et al. 2012). The most successful quantitative meeting of theory and fact within the fitness-maximisation paradigm is in the study of sex ratio. The recent book by West (2009) is now the standard work, and readers who may need persuading of the astonishing power and empirical success of the paradigm are urged to consult it. The subject of Animal Behaviour, or Ethology, has suffered turmoil over the increasing dominance of design questions, with many practitioners pleading for retention of plurality of interest by continuing to study also the other three of the 'four whys' that Tinbergen (1963) offered as defining the biological approach within science: namely not only *function* but also *ontogeny*, *phylogeny* and *mechanism*.

The design approach told us most of what we know about natural selection, via Darwin, and fitness-maximisation has been an astoundingly successful approach at uncovering biological facts and patterns since the revolution of the 1960s and 1970s that unleashed evolutionary ideas on whole organism biology. First-year biology undergraduates are taught that natural selection results in organisms acting as if maximising their fitness, researchers apply for grants on that assumption, and funding agencies award many grants taking fitness-maximisation as a given.

A fly in this ointment is that there are serious reasons to doubt that fitness is in fact maximised. The central assumption of the approach has been known to be untrue in general for decades, and it is here that the other of the two approaches to studying natural selection becomes relevant. Following the reconciliation between Darwinism and Mendelism brokered by Fisher (1918), biologists agreed in the Modern Synthesis that evolution operates through changes in gene frequencies. They construct population genetic models by defining a set of genotypes, and tracking the changes in genotype frequencies (in principle for each age and sex) through a life cycle that can include birth, survival to breeding, mating fertility, survival for further breeding etc, and then death. It is these models that should tell biologists whether 'fitness-maximisation' is a well-founded principle, and the resounding answer has been 'no' (Ewens 2004).

There have been a number of attempts to tackle this apparent conflict. The earliest formal claim for fitness-maximisation was Fisher (1930)'s fundamental theorem of natural selection, whose original (notoriously imprecise and misleading) verbal version reads "The rate of increase in fitness of any organism at any time is equal to its genetic variance in fitness at that time". Formal demonstrations of fitness-maximisation, of which the fundamental theorem was the first, can be regarded as providing a 'licensed anthropomorphism' (Grafen 2003) in place of a naive and unreflective one. This theorem inspired the adaptive landscapes of Wright (1969–1978), which are still much used in an informal way to explore the consequences of assuming fitness-maximisation in evolution. An adaptive landscape imagines an x-axis or set of x-axes representing genotypes or phenotypes (the choice is controversial) with the fitness of each raised on a y-axis, representing the 'height' of the landscape. The theorem itself has a long and difficult history: it moved from initial acceptance following its republication in Fisher (1958), to a

chorus of rejection following Moran (1964)'s famous paper 'On the non-existence of adaptive topographies', a reinterpretation and restoration of Fisher's argument and conclusion by Price (1972), the rediscovery of Price's work by Ewens (1989) and Edwards (1994), and its current acceptance as a true theorem but with no obvious meaning (see Ewens 2004, for an authoritative population genetic view). The tortuous history of this first formal claim illustrates the misunderstandings between population geneticists and design-oriented biologists over what is meant by fitness-maximisation.

One of the admirers of the fundamental theorem was Hamilton (1964), who modelled his theory of inclusive fitness on Fisher's theorem, explicitly aiming for and achieving a parallel fitness maximisation principle. This generalised the fundamental theorem, and provided a more sophisticated quantity to stand as fitness when social interactions were permitted in the underlying model. Fisher had explicitly assumed these away.

The renaissance of fitness-maximisation ideas among biologists in the 1960s and 1970s did lead some biologists, most notably Lloyd (1977) and Maynard Smith (1978), to investigate its central tenet through the study of simple theoretical examples. A telegraphic version of their conclusions is that apart from exceptional cases, it is reasonable to proceed on the basis that fitness is maximised. Here are some exceptional cases, to convey a sense of the technical challenge in maintaining fitness-maximisation.

1. Overdominance. Suppose a single locus has two alleles, and that the heterozygote AB has a higher survival s_{AB} than both homozygotes, s_{AA} and s_{BB}. Then there is an equilibrium frequency of A given by $p_A^* = \frac{s_{AB}-s_{BB}}{2s_{AB}-s_{AA}-s_{BB}}$. At this equilibrium, natural selection may be regarded as trying but failing to increase the fraction of heterozygotes; it fails because recombination creates half heterozygotes and half homozygotes in a brood of offspring even when both parents have the optimal heterozygote genotype. Clearly, the fittest phenotype is not the only phenotype seen at equilibrium, and furthermore it would be impossible to understand the distribution of phenotypes without knowing the genetic system underlying them.
2. All heterozygotes. In the same system, suppose we had a population consisting of all heterozygotes AB. Then one generation later, the distribution would by Mendelian genetics be a quarter each of the homozygotes AA and BB, and a half would be heterozyotes AB. The mean fitness would have decreased. Even once a population consists of all optimal phenotypes, therefore, it can, for reasons that are entirely clear if the genetic system is known, move far away from that state.
3. Separate sexes. Suppose the fitnesses are reversed, specifically $s_{AA} = s_{BB} > s_{AB}$, and that the population consists of AA males and BB females. The population is currently optimal in the sense that each individual has an optimal phenotype, but we know that next generation all individuals will be AB and suboptimal.

In the face of these examples, the simple view is that there is no useful fitness-maximisation principle available. Moran's famous examples involved two loci, but for current purposes these single-locus cases will do.

In 1984, I coined the term 'Phenotypic Gambit' for the research strategy of studying organisms in ignorance of the actual genetic architecture of the trait in question, and further as if the simplest architecture applied, namely a haploid genetics in which there was a separate allele for each element in the set of possible phenotypes. As well as illustrating with examples that this research strategy was in fact much used and very productive, I gave some reasons to think it reasonable, but also insisted that 'we should recognize the urgency of the need to provide a proper justification for this convenient simplifying assumption' (Grafen 1984, p65).

It is helpful to keep in mind three kinds of players of the phenotypic gambit:

1. Field-worker. McGregor et al. (1981) studied song repertory size in male great tits in Wytham Wood in relation to their reproductive success.
2. Modeller. Maynard Smith and Price (1973) constructed a simple game theory model to investigate the natural selection of aggression and its limitation.
3. Comparer. Ridley (1983) used cross-species data to investigate the hypothesis that assortative mating for size evolved because (i) larger females are more fecund and (ii) larger males tend to win fights over the more fecund females.

These examples could be multiplied thousands-fold from the literature. The Field-Worker could nowadays obtain some empirical information about genetic architecture fairly easily, as discussed in Sect. 5, but is that necessary? The Modeller is dealing with no particular species, but wishes the results to apply widely. The Comparer's whole exercise depends on the phenotypic gambit—if it really is necessary to do genetics, then a genetic exercise for each datapoint would be required: and then, it is hard to see what the value of the study is.

We can imagine a range of possibilities. At one extreme, genetics makes no important difference at a phenotypic level, and the Phenotypic Gambit is entirely justified. At the opposite extreme, genetics always makes a considerable difference and cannot be ignored. In this case, all previous studies are worthless except as hints for possible properly genetically informed studies. Our apparent success in understanding the design of a kidney, of an eye, or a liver, would have to be regarded as illusory. Camouflage patterns would after all be incomprehensible except in terms of their yet-to-be-discovered genetic architecture. The substantial body of work on theory and empirics of mimicry, including its division into Batesian and Mullerian subcategories, would need to be completely reworked. The patterns of sex ratio reviewed by West (2009) would have to be reinterpreted as mistaken. It is clear that a great deal of biology as currently researched and taught would simply not have seriously begun. But it is also clear that if we are persuaded by any of those large significant areas of biology in their current state, then we reject this extreme possibility. Attention then focusses on the messier but more realistic questions of how much difference genetics makes, and this would likely vary from area to area. The Phenotypic Gambit articulates the assumption that is usually made implicitly in this work, and the formal darwinism project aims to understand better why and how the gambit works when it does, and also to identify and understand those cases in which the gambit fails.

An intermediate possibility is that, in some rough metric, the Phenotypic Gambit gives results that are 10% out. It seems likely that most of the biology would still be

worth doing in cases where the genetics couldn't be done; and indeed, whether the genetics would be worthwhile discovering would rather depend. A comparative study with 60 species would need a lot of genetics done: it seems likely that the genetics would simply add complication and noise to the biology, rather than anything of biological interest; and it also seems likely that geneticists would not be especially interested in uncovering the genetics of a trait chosen for its biological interest. Thus, once on the realistic quantitative slope of "how much difference does the genetics actually make?", it seems quite likely that the sensible response would be 'while the errors due to genetics are thought to be tolerable, keep going without the genetics; if the genetics is thought essential, then abandon the study altogether as the extra costs and complications are not worthwhile'. There are many conjectures in this position, and many biologists would disagree with them and with my conclusion. Material with which to have that discussion substantively is increasingly available, with the recent accumulation of empirical evidence about the genetics of biological traits, and especially in microbes. This literature is discussed in Sect. 5.

It is worthwhile at this point to consider *who* is assumed to be doing the 'fitness-maximisation'. Much biology since 1955 can be arranged around this question (Segestråle 2000), which encompasses questions of levels of selection and the individual versus group versus species selection debates. The approach taken in the formal darwinism project is that it is the individual that has a fitness and that may or may not be maximising it. We see this in a number of ways. First, the set of possible phenotypes is the set of phenotypes that an individual may possess—not the set of phenotypes that a group or species may possess, and not the set of phenotypes that a gene may possess. (For example, the phenotype of a group would surely contain information about the number of members, their ages and sexes, as well as details about each individual.) Technically, then, we are dealing with the maximisation of individual fitness. Correspondingly, the evaluation of a phenotype is the number of offspring an individual with that phenotype has. Gardner and Grafen (2009) employ the methods of the project to investigate other possible candidates for who does the maximising.

The formal darwinism project was my own, somewhat belated, response to my 'urgency' of 1984, begun when I saw what shape such a response could take. A prospectus for a proper justification would include a number of generalities. It should apply to a wide range of genetic architectures, not depending on the number of loci or alleles involved, and if possible accommodating epistasis, dominance and other genetic complications. It should where possible not depend on knowledge of the mating system. These generalities serve two sets of purposes: first, they would provide a reasonable justification for the Field-Worker, the Modeller and the Comparer, who will rarely know those extra details; and second, they also provide a retrospective justification for Darwin's conclusions, by showing that the view that natural selection could create design was robust against genetic details which Darwin had no choice but to ignore. Thus, formally an abstract connection between the design approach and population genetics, the project's biological significance stems from both its defence of current behavioural ecological (and other design-based) practice, and its exploration of the logic of Darwin's original argument.

3 Content

This section will describe the links established between the fitness-maximisation and population genetic approaches, but leave most of the interpretation and implications to the next section. We begin with the formal structures used to represent each of the approaches. Population genetic models (Ewens 2004) are examples of dynamical systems, either differential or difference equations, and are usually highly specific about the genetic architecture. The covariance selection mathematics of Price (1970,—specifically not the extension of Price (1972)) does offer just the generality set out in the prospectus of the previous section. To represent the design approach, it is natural to take optimisation programs, as used in microeconomics, operations research and game theory, as the formal structure. The project then became about constructing formal links between the mathematics of motion and the mathematics of optimisation.

The Price equation is very slippery. By itself merely an accounting identity, the Price equation is powerfully suggestive, often too powerful. Let us follow an oversimple argument purporting to link optimisation with gene frequency change. Positively, some features associated with the Price equation are shared with the formal darwinism project. Negatively, the gaps in the argument show where the project has to step carefully, sometimes at considerable length that will not be repeated here.

The classic formulation of the Price equation has a parental population I, and assumes we know the number of successful gametes w_i for $i \in I$. The focus of attention is the change in a p-score, which can be just the frequency of a single allele, but can also be an arbitrary weighted sum of allele frequencies. A novelty of Price's approach is that this gene frequency is defined for individuals, and we assume we know p_i for $i \in I$. If a simple allele frequency, then p_i is zero or one for haploids, and may also take the value a half for diploids. We further assume we know Δp_i, the difference between the average p-score of the successful gametes of i and i's own p-score. Notating averages by dropping the subscript, using $\mathbb{E}$ and $\mathbb{C}$ for expectation and covariance (in the sense of population averages), and Δp for the change in mean p-score from one generation to the next, the Price equation states that

$$\Delta p = \mathbb{C}\left[p_i, \frac{w_i}{w}\right] + \mathbb{E}\left[\frac{w_i}{w}\Delta p_i\right]$$

Three significant points may be made straight away about this equation. First, to calculate the right hand side requires a great deal of knowledge about the individual parents, their genotypes and their successful gametes, while the left hand side provides only the change in mean p-score between the generations. Thus it does not provide enough information to crank the handle again and move on to the next generation. It is, however, the assumptions of the Price equation, rather than the equation itself, which should be termed 'dynamically insufficient', for we have made no assumption about how the successful gametes are combined to form the individuals of the offspring generation (Frank 1995). Second, it holds with an extreme generality, irrespective of practically all details of genetic architecture and

mating system—except that all the alleles involved in the weighted sum whose change is calculated must belong to the same coreplicon (Cosmides and Tooby 1981), which for mammals means either that all the alleles are at autosomal loci, or that they are all on the X-chromosome, or that they are all on the Y-chromosome. Third, it suits a project of linking to the design approach because it indexes individuals and so focusses on individuals, rather than on genotypes: indeed, it holds in exactly the same form simultaneously for all p-scores, including therefore the frequency of every allele in the coreplicon. (As an aside, the choice of coreplicon is fundamental. If we choose X-chromosomes in humans, then when we count the number of offspring of a man, we count only his daughters. In generality, choice of coreplicon affects even the values of w_i.)

Let us look at the obvious informal argument linking the Price equation to optimisation. Suppose the second term on the right hand side is small enough to be ignored, which will usually be the case under Mendelian segregation. The frequency of any allele that positively covaries with w_i increases, and the frequency of any allele that negatively covaries with w_i decreases. If a trait is under genetic control, it seems plausible that these changes in gene frequency will succeed in changing the trait in such a way as to increase overall the values of w_i in the next generation. If that is true, then the changes in gene frequencies increase w, and so fitness is increased. Indeed a slightly formalised version of this argument says that every trait has an additive genetic value, which is an allelic weighting that best predicts the trait (those with statistical training may like to know that 'best' is in the sense of least squares). Let the additive genetic value of fitness itself be a_i, so that $w_i = a_i + \epsilon_i$, where ϵ_i are residuals that sum to zero and are exactly uncorrelated with a_i. Noting that a_i is a weighted sum of allele frequencies, and so a p-score, one line suffices to show that

$$\Delta a = \mathbb{C}\left[a_i, \frac{a_i + \epsilon_i}{a}\right] = \mathbb{C}\left[a_i, \frac{a_i}{a}\right] = \frac{\mathbb{V}[a_i]}{a}$$

and so the change in the mean of a_i equals the variance in a_i divided by the mean of a_i. Some readers may recognise this as a version of Fisher's Fundamental Theorem (though so highly simplified that it omits the age-structure, and therefore also the survival-fecundity tradeoffs, represented in the original). Nevertheless, this simple form does seem to support the view that mean fitness is systematically increased (i.e. $\Delta a > 0$) while fitness is heritably non-constant (i.e. $\mathbb{V}[a_i] > 0$), and so that fitness is maximised under natural selection.

This plausible argument fails—one key point is that fitness and its additive genetic value have a changing and complicated relationship over time. Rather than discuss the deeper principles, it suffices here to see that the difficult examples of the previous section are situations in which: (1) the mean fitness could increase but doesn't; (2 & 3) the mean fitness decreases.

A further point that exhibits the characteristic slipperiness of the Price equation is that this argument, if valid, would prove too much in the sense that it would imply false conclusions. Without restricting how w_i arises, the Price equation remains true even if there are social interactions, and social interactions don't immediately seem to affect the rest of the argument. Yet believers in fitness-maximisation themselves

(claim to) know that when there is social behaviour, it is inclusive fitness that is maximised, rather than classical fitness. There are cases in which the increase in inclusive fitness brings about a reduction in classical fitness, and maximisation of one implies non-maximisation of the other.

Careful pinning down is necessary to construct demonstrably valid arguments when there are so many plausible but invalid arguments. For example, Grafen (2002) constructs a formal model making w_i an arbitrary function of the phenotype of individual i and chance events, excluding all other dependencies, while Grafen (2006a) handles social behaviour by allowing w_i to depend on the phenotypes of all individuals, not just i. Here we leave the population genetic side (just as things are getting interesting, some readers may feel, but they are also getting too technical), and move on to how the design approach is formalised.

Pinning down of the Price equation is thus necessary because the equation itself is so suggestive, encouraging readers to embrace untrue general conclusions. The main pin is the formalisation of the design approach using an optimisation program. Most treatments including the original (Fisher 1930) content themselves with informality on the optimisation side, and the major new technical feature of the formal darwinism project consists in setting up optimisation programs explicitly, and proving formal links to them.

The optimisation program side is simpler in itself. Because population genetics is fundamental, the first task is to construct the optimisation program from the contents of the population genetic assumptions. An optimisation program contains three elements. One variable, the *instrument*, is regarded as under the control of the implicit optimiser, who selects a value of the variable from some set of possible values, the *constraint set* or *feasible set*. A real-valued function defined on that set, the *maximand*, indicates the degree of success of each possible choice. A choice of instrument is said to *solve* the program if it achieves the highest possible value of the maximand within the feasible set. In our application, the instrument is simply the phenotype, and the project assumes that the set of possible phenotypes is specified on the population genetic side—it comes into play there, for example, in considering mutants that test an equilibrium. Thus the instrument and constraint set are simply carried over from population genetics, but there remains the question of how to construct the maximand. It is notorious that biologists who "know" that fitness is maximised can frequently be found discussing what exactly fitness is; and this step in the project corresponds to answering that question.

So far in the formal darwinism project, in both the two fully developed cases (Grafen 2002, 2006a), the maximand has been proposed by informed guesswork, and then the guess has been justified by proving links between the optimisation program so constructed and the Price equation.The links aim to show that the maximand fulfils the biologist's concept of fitness; they are made strong to render it is unlikely there could be any other choice for maximand that would also fulfil the links. An individual will be said to 'solve the optimisation program' if its phenotype achieves the highest feasible value of the maximand. The links differ in detail in the different cases of the project. Here are schematic versions of them:

1. If all individuals in the population solve the optimisation program, then the expected change in every gene frequency equals zero, and there is no possible phenotype which, if produced by a rare dominant mutant, would initially invade the population.
2. If all individuals attain the same value of the maximand but do not solve the optimisation program, then the expected change in every gene frequency equals zero, but there is a possible phenotype which, if produced by a rare dominant mutant, would initially invade the population.
3. If individuals attain different values of the maximand, then the change in every gene frequency equals its covariance across individuals with those attained values.
4. If the expected change in every gene frequency equals zero, and if there is no possible phenotype which, if produced by a rare dominant mutant, would initially invade the population, then every individual in the population solves the optimisation program.

The first three make assumptions about the optimisation program, and draw conclusions about gene frequency change, while the fourth moves in the opposite direction. Taken together, they show that the maximand embodies most of what a biologist would wish 'fitness' to mean. Note that adding a constant to the maximand retains all of these properties, but multiplying by a positive constant does not (one side of Link 3 would scale with the factor, but the other side would not). Other considerations (Grafen 2007b) show that zero is a meaningful value of fitness, and so all together fitness is uniquely defined.

One important task is to see how they cope with the three difficult examples of the previous section. A link will be said to hold 'trivially' if its hypothesis ("antecedent" for philosophers) fails. For example, Link 1 says if all individuals solve the optimisation program, then such and such follows. With overdominance, not all individuals solve the optimisation program, so Link 1 does hold there, but in a trivial sense because the exact nature of 'such and such' is irrelevant.

1. Overdominance. Links 1 and 2 hold trivially. Link 3 applies, though at equilibrium it turns out that the covariance of allele frequency with fitness equals zero. Link 4 holds trivially because the second hypothesis fails: if a dominant mutant produced the 'heterozygote' phenotype, it would initially invade.
2. All heterozygotes. Link 1 holds substantively, though it is important to note that the conclusion is that *gene* frequencies do not change (indeed they remain at 0.5), even though *genotype* frequencies do change. Links 2 and 3 hold trivially. Link 4 holds substantively—indeed, every individual in the population solves the optimisation program.
3. Separate sexes. The same pattern just stated holds, though in Link 4 we note that here the individuals solve the optimisation program with two different genotypes, which could in principle also be two distinct phenotypes.

Thus the links survive, which as mathematically proven results they must, when tested against the difficult cases. The way in which they survive may, however,

seem to contain an element of evasion, and call into question the meaning and value of the links themselves.

This section now ends with a brief description of the three main core papers so far in the formal darwinism project. Grafen (2002) proves links of the kind just described, on the assumption that number of offspring depends only on an individual's own phenotype and on chance events. One main conclusion is that fitness must be defined relatively, that is, as number of offspring divided by the population mean number of offspring in any particular environmental outcome, and another is that the maximand is the probability-weighted arithmetic average over uncertainty of that relative number of offspring. The population and the set of environmental outcomes may be finite or infinite. Grafen (2006a) also proves links of the same kind, but permits social interactions, with a finite population and finite set of environmental outcomes. It requires that number of offspring should depend additively on the phenotypes of other individuals. The maximand is the probability-weighted arithmetic average of relative inclusive fitness, where inclusive fitness is the usual sum of effects on fitnesses weighted by relatednesses, but it is relative to the mean neighbour-modulated fitness, *not* to the mean inclusive fitness. Relatedness is defined in the course of the paper. These are the two fully-developed cases.

Grafen (2006b) is the third, not fully developed, case. It accepts that offspring may not all be equal, and instead allows them (and therefore their parents too) to belong to a class in some set. The set of classes in mathematical terms could for example be a compact subset of $\mathbb{R}^n$: biologically, class could be defined as depending simultaneously on one or more continuous variables such as height or tarsus-length, and also on discrete variables such as sex and number of vertebrae. The paper shows how to assign relative weightings to the classes (essentially generalising Fisher (1930)'s approaches for age and sex), and its main result is a reproductive-value-weighted Price equation. The change in mean p-score is weighted over the classes, and the fitness of each individual parent weights offspring, according to the same reproductive values. The notation and arguments needed to achieve this equation were so involved that it was not possible in that work to prove versions of the usual set of links, and nor was it possible to incorporate uncertainty, though some links were proved in special cases. All three papers assume discrete generations, though overlapping generations are finessable through the usual device of considering a surviving parent as a special kind of offspring. In all three papers, the genetics is arbitrary in the usual way for the Price equation.

With this schematic characterisation of the content of the results of the formal darwinism project, we move on to consider in the next section the implications of that content.

4 Implications

The way the typical links hold for the three difficult cases may seem somewhat evasive. On the other hand, the needs of the Field-Worker, Modeller and Comparer could be well met by links of that kind if they held in a significant way. The aim of

this section is to explore what the links can tell us about the operation of natural selection.

The first point is that, although not formally considered as yet, it seems likely that the requirements for the four links to be met are so strong that changing the maximand (other than by adding a constant) would break the links. This would imply that the candidate maximand found is essentially the only possible maximand, which is significant because other kinds of candidates have been considered in biology including, for grouped populations, the group's number of offspring.

Second, the candidate maximand has special properties that are worth knowing. For one thing, fitness is relative to the population mean, and so in the 2002 case its mean is always 1. Perhaps a mere curiosity, with social interactions, inclusive fitness has to be made relative to the mean neighbour-modulated fitness, and so its mean will generically be less than 1 (owing to the 'diluting factor' of Hamilton 1964). Relativity becomes critical when there are different population sizes in different environmental outcomes: to give a crude example, one offspring in a population of 1000 is worth more than three offspring in a population of 5000, showing that the mean and variance of absolute number of offspring can have only indirect connections to gene frequencies in the next generation. Further, the maximand is an arithmetic mean over uncertainty: specifically not a geometric mean, nor some other aggregating statistic that would imply a mean-variance tradeoff. When there are classes, the weightings for different offspring classes are the same for each parental class. These properties are given extra significance by the (admittedly tentative) uniqueness property of the previous paragraph—there is no escape to some other notion of fitness that might be maximised instead. It is the candidate or nothing.

The generalities of the argument provide the third significant point. It is the same candidate maximand that applies for all the genetic architectures. Although the holy grail of proving that fitness-maximisation is always achieved in theory and practice is not attained, and obviously never can be as it is not always true (as shown by the difficult examples of Sect. 2), it is important that we almost reach it with a single argument for all genetic architectures. This opens the possibility, not yet pursued, of using formal means to classify the exceptional cases in which fitness-maximisation does not occur. For the moment, we at least see that there is a very general expectation of something close to fitness maximisation, which will convert into fitness-maximisation unless there are particular kinds of circumstances—and further, that fitness is the same quantity for all genetic architectures.

The final implication to be discussed of the results of the project so far is the uncovering of assumptions. One key point in the linking argument occurs when the population in the population genetic model is replaced with a single decision-taker in the optimisation program. To take this step requires assuming all individuals are essentially the same from a strategic point of view, and in the 2002 version I introduced *pairwise exchangeability*. "One relevance of this is to ensure that we do not have on average an association between, for example, one allele and particularly favourable environments. The major reason is to ensure that all individuals face the same environmental challenges, and so are having to solve the same problems." (Grafen 2002). Discovering previously unrecognised assumptions—and especially

being confident there are no others—is an important and valuable consequence of constructing formal arguments.

We now consider the relationship between the formal darwinism project and adaptationism, which I take to be closely allied to the Phenotypic Gambit of Sect. 2, namely the view that it is productive to regard organisms as well-designed in relation to their environment. First, the project explores making adaptationism more precise: setting up an optimisation program formalises adaptationism, and by using population genetics to forge links, the project says, for example, 'If organisms are well-designed through natural selection, then the maximand must be of such and such a form'. The project also aims to allow a classification of the exceptional cases, in which natural selection does not lead to fitness-maximisation. All of these points are neutral about the truth of adaptationism. However, it is admitted that one motivation behind the project is indeed to defend a carefully delineated version of adaptationism, against a variety of mathematical and other objections to it, particularly because many important areas of biology rely on adaptationism, as discussed in Sect. 2.

To these general points can be added two further sets of implications of the project. The assumptions made in versions of the argument are very general, and so the results may be applied to other models. The first set arises from this use of the core theory as a meta-model. Two examples may be mentioned. Killingback et al. (2006) published a model of altruism in a population comprising groups of different sizes, and claimed that this model exhibited altruism that could not be attributed to inclusive fitness. This model met the assumptions of Grafen (2006a), and I subsequently showed (Grafen 2007a) that the altruism produced was precisely attributable to inclusive fitness. Ohtsuki et al. (2006) published a network model and, again, claimed that altruism evolved because of forces other than inclusive fitness. Again this model met the assumptions of Grafen (2006a), and I was able to show (Grafen 2007c) that the altruism was precisely attributable to inclusive fitness. As the project embraces more situations with more generality, it will be able to assist with the interpretation of a wider and wider set of models.

The second set of implications stems from viewing the project as articulating Darwin's central argument in *The Origin*, and also the central arguments of later additions to darwinism, as they relate to design. The work of Gardner and Grafen (2009) applying the project's methodology to distinguish 'group selection' and 'group adaptation', and to investigate the latter, has already been mentioned. Turning to non-biologists' interventions, it is notorious that new interpretations of Darwin frequently appear, in a way that does not happen for scientific work with a firm mathematical framework. If the formal darwinism project is accepted as articulating modern biology's understanding and updating of Darwin's central argument, then this would set the bar higher for physicists and others who wish to set aside current orthodoxy and propose their own new Darwinism.

5 Connected work

Rather like some branches of philosophy, the formal darwinism project leaves most things where they were. For most behavioural ecology work, it may provide a

background reassurance that the genetic details, which aren't in any event known, are unlikely to matter. Theory about how gene frequencies change will continue unaltered. The literatures to be discussed briefly in this section are those that are frequently cited as justifying fitness-maximisation, theory that aims to justify inclusive fitness or Darwinian fitness, theory that assumes fitness-maximisation, theory that appears to contradict the claim that the arithmetic mean over environmental uncertainty is maximised, with no regard to the variance in fitness; and work that aims to discover the genetic architecture of 'real' traits.

Users of inclusive fitness often cite Taylor (1990), (1996), who has already brought together inclusive fitness and a finite number of classes to which individuals belong, and shown how reproductive value illuminates the operation of natural selection. However, he really offers a formalism, and does not attempt to prove results, so the boundaries within which his papers consider inclusive fitness are not altogether clear. His focus is on whether inclusive fitness calculations correctly predict selection at a single locus on the assumption that that locus is the only determinant of phenotype, and he does not consider an arbitrary set of possible phenotypes. Fitness-maximisation is not explicitly part of the agenda in these papers, despite the use to which many citing papers put them. Rousset (2004, and references therein) provides powerful tools and analytic results about social selection in usually symmetrically subdivided populations. Like Taylor, he focusses on selection at a single locus with two alleles, and does not consider phenotype spaces and fitness-maximisation, although is also cited to support the use of inclusive fitness maximisation. Taylor and Frank (1996) and Frank (1998) offer excellent and highly influential advice on how to construct a kin selection model, but this is 'downwards looking', towards applications, and does not prove results about fitness-maximisation in general. Hammerstein (1996) proposes a 'streetcar theory' that aims to connect fitness-maximisation with gene frequency change in the presence of frequency-dependence, and considers destinations of evolution: a biologist's concept of fitness also says a lot about selection in progress. Multi-locus methodology (e.g. Kirkpatrick et al. 2002; e.g. Gardner et al. 2007) has similar goals of generality about genetic architecture, but does not focus on fitness-maximisation. Retaining dynamic sufficiency means retaining a large formal superstructure for the genetics, and it remains to be seen whether this can be managed at the same time as the complexities introduced, for example, by sophisticated phenotypes and linking to optimisation. These papers inhabit the borderline between the use of fitness in population genetics and fitness-maximisation, and have important implications for their relationship, but without explicitly tackling the link between them.

Two literatures that assume fitness-maximisation are now considered. 'Adaptive Dynamics' (e.g. Meszéna et al. 2001) makes simplifying assumptions so as to be able to assume fitness-maximisation, which is then exploited at a higher level to study speciation and longer-term evolutionary change. Houston & Mc-Namara (e.g. 1999) present a pioneering set of models for particular applications of a kind that are highly relevant to biologists. An important goal of the formal Darwinism project is to contain as many as possible of these models as special cases, which gives a very concrete and biologically significant 'target' for the abstraction. The formal

Darwinism project is precisely about proving links between population genetics and fitness-maximisation in the first place. Work in this area does explore what the fitness being maximised should be, and so in principle the formal darwinism project could at some stage assist.

I now turn to the idea of bet-hedging. Frank and Slatkin (1990) review and systematise the theory about natural selection under environmental uncertainty from Dempster (1955) onwards. They begin with a 'general model' in their equation (3) that is very similar to the Price equation in the formal darwinism project, though not quite so general. They present a generalisation of some but not all previous work that the allele more likely to fixate is the allele with the highest value of $\mu_i - q_i \rho_i \sigma_i^2$, where μ_i and σ_i^2 are the mean and variance of the probability distribution of the absolute number of offspring of individuals with allele i, q_i is its frequency, and ρ_i is the correlation in absolute numbers of offspring between individuals bearing that allele. They propose a key quantity $k_i = q_i \rho_i$, which is the correlation between an individual's reproductive success and the average reproductive success of the population, under what I will call the "no correlation assumption", namely that there is no correlation, across the range of environmental uncertainty, in absolute fitness, between individuals with two different genotypes. The result $\mu_i - k_i \sigma_i^2$ neatly encapsulates the idea, widespread in biology, that there is a mean-variance tradeoff in determining which phenotype we should expect to see at equilibrium under uncertainty. Why does this tradeoff not appear in the formal darwinism project? Frank and Slatkin's aim was to unite and generalise previous results that were expressed in terms of the means and variances of absolute fitnesses and concluded which of two alleles is likely to spread. In constructing abstract links between gene frequency change and individual fitness-maximisation, more generality can be obtained by working with relative fitnesses. Frank and Slatkin have mean-variance tradeoffs because they work with absolute fitnesses, while the formal darwinism project has no mean-variance tradeoff because it deals only in relative fitnesses.

Thus which kind of analysis is regarded as distilling the essence of the effect of uncertainty on natural selection is a matter of purpose. My view is that relative fitness provides greater generality. The particular examples of mean-variance tradeoffs Frank and Slatkin (1990) use to illustrate their theory rely on the "no correlation assumption", and under diploidy there are further assumptions effectively including random mating. It is also relevant to point out that the maximand for individual fitness-maximisation is always the same in the formal darwinism project, in the sense that it does not depend on further details of the model, while the precise tradeoff between mean and variance (for example, whether $k_i = 1/2$ or k_i equals the inverse of the population size) does depend on them. In these senses, the maximisation principle being investigated by the formal darwinism project is considerably more general. Of course it may be that further work could extend the range of the mean-variance tradeoffs demonstrated by Frank and Slatkin (1990).

Work by Lande and collaborators (e.g. Lande 2007, e.g. 2008, e.g. Engen et al. 2009, 1981) extends understanding of uncertainty in demography and selection in an age-structured population, using a quantitative genetics framework. They also

express their results by saying that, under their approximations, the adaptive topography in the presence of environmental uncertainty has as its ordinate the mean minus half the variance of offspring number, because they also deal with absolute number of offspring. Maximising the mean minus half the variance of absolute offspring number is an approximation that is sometimes equivalent to maximising the arithmetic mean of relative number (Grafen 1999). The results of Lande and colleagues are obviously not under any question: they are making approximations to find linear and quadratic expressions under weak selection in a quantitative genetics framework, while the formal darwinism project seeks a natural way to understand the evolutionary response to uncertain environments in a more abstract setting. There is thus no contradiction between the two approaches.

On a methodological note, Ellner and Rees (2006) have used similar advanced mathematical techniques, involving invariant measures of Markov processes over abstract topological spaces, to Grafen (2006b), also to study reproductive value, but in an ecological and demographic, not evolutionary, study.

One major purpose of the formal darwinism project is to assist in showing that the exact genetic architecture may reasonably be ignored in many circumstances when the form of a trait is the focus of biological study. In 1984, I wrote:

> Genetically simple and well-studied characters are rarely of evolutionary interest. They are usually straightforwardly disadvantageous mutants maintained by judicious artificial selection in strains which have spent tens of generations in the laboratory. (Grafen 1984, page 65).

Over the past three decades, staggering advances in molecular biology have revolutionised what has been and can be discovered about the genetics of real traits in natural populations. Dr Jarrod Hadfield works in this area and I quote with his permission from a communication to me: "Barrett et al. (2008), for example, has gone a long way to understand the molecular genetics of an ecological trait, but this is still rare. If you exclude simple Mendelian traits like coat colour in Pocket Mice, horn type in Soay Sheep and colour polymorphisms in Gouldian Finches and Snow Geese then we know very little about the molecular genetic basis of most traits. The majority seem to be so polygenic and the effect of each locus so small that identified polymorphisms usually explain a tiny fraction of the genetic variance (for example see Manolio et al. 2009). More broad scale molecular genetic analyses looking at the distribution of (for example) dominance coefficients would be really useful but still very difficult - not only do we not know which are the causal polymorphisms we usually haven't even typed them and are working with polymorphisms that are in linkage disequilibrium with the causal polymorphisms." With bacteria, it is another matter, where the genetics is more straightforward to study but the natural ecology is usually mysterious, and that interface is reviewed by West et al. (2006). Combining genetics and ecology is an active area in which surprises are likely at both these scales.

It is fundamental biology to know the genetics of traits, and obviously interesting to link that to adaptation, so these kinds of study clearly have their own independent interest and value. Nevertheless, particular questions *are* raised by taking this area and the formal darwinism project together: I will mention some of them, even

though answers are beyond the scope of this paper. One is whether they are discovering further examples of over-dominance, or other similar types of architecture, in which the distribution of types is wholly incomprehensible without knowing the genetics. Another is whether it is possible to formulate in some way how inaccurate non-genetic study is likely to be, on average, when a trait is studied for its adaptive interest. These two questions continue to press the issue of whether it is reasonable to ignore genetics if one's interest is phenotypic. A third question has a different slant: how often are traits under the influence of more than one coreplicon? This matters because the whole theory of the formal darwinism project so far operates on the basis that a trait is under the control of a single coreplicon: traits that are often under joint control of, for example, autosomes and sex chromosomes, would pose special puzzles. I am grateful to Dr Hadfield for pointing me to the work of Postma et al. (2011) on shared control of guppy coloration of just such a kind, who note that shared control is likely in the presence of 'sexually antagonistic selection' (van Doorn 2009) in which selection acts at the same locus in opposite directions in the two sexes.

The formal darwinism project is thus by no means operating in a vacuum and, as the core theory develops, its capacity increases for conceptual engagement with other literatures.

6 Discussion

Two points will be covered in this discussion: further work on the core theory of the formal darwinism project that is in the pipeline, and why this paper was invited for publication in a journal linking Philosophy with Biology. Dr Richard Gratwick and Dr Paul Crewe have been reworking the core theory of the formal darwinism project in a more rigorous and mathematical way. Although the core papers already published seem highly mathematical to biologists, they are very unmathematical to mathematicians. One point of the reworking is to make sure that at least one audience will find the papers comprehensible. A second is that further developments are much easier from a systematic and logically set out base. It has already been mentioned that the 2006 paper on classes and reproductive value did not manage to prove formal links to optimisation. The new framework begins with a manuscript (Batty et al., in press) in which a result embracing classes and reproductive value as well as uncertainty, corrects, unites, and simultaneously generalises two earlier papers (Grafen 2002, 2006b). The great virtue of formalisation is that new steps become possible: one radically new argument extends fitness-maximisation ideas to demographic as well as environmental uncertainty, and another provides a simple route to abstract inclusive fitness results. Thus although the general philosophy of linking population genetics to fitness-maximisation remains, the core theory of the formal darwinism project is undergoing rapid advances.

Ideally, a single result would encompass inclusive fitness, environmental and demographic uncertainty, and discrete and continuous time. Darwin did not concern himself with these details while establishing most of the general truths now known

about natural selection. A modern biologist should be able to rely on formal results that achieve the same kind of concentration on the biological essentials.

Turning towards philosophy, possibly wider implications can be introduced by looking at the optimisation programs as further articulated in Sect. 5 of Grafen (2002). Agency is the idea that a person is not merely a machine, but has choices to make about acting in the world. To the extent that the idea of agency is motivated by introspection, there is nothing further to be said here. To the extent that agency is inferred or imputed because of complex and purposive behaviour, then the project could be relevant. The formal work regards an organism as making one decision in the face of uncertainty, and also possibly as in possession of partial resolution of that uncertainty. This single decision may be complicated because it represents the exact shape of some body part that is definitely produced. On the other hand, the formal arguments do not distinguish that case from one in which, while formally a single decision, the organism is actually making a whole series of conditional decisions over time in informationally-complicated situations. Grafen's development shows that optimality in such a program immediately implies acting as if in possession of a correct prior distribution over the uncertainty, and also as if performing optimal Bayesian updating of that prior in the light of the partial information, at each stage of the decision-making process. The formal arguments also say nothing about the mechanisms by which these decisions are implemented, and for rapid decisions with much information and complex calculations required, a brain would be an effective mechanism. Thus, if agency is imputed because of elaborate goal-directed behaviour, then the project may be connected to agency through predicting just such behaviour, and supplying the sketch of an outline mechanism for its evolution by natural selection. An organism solving the program therefore seems to be equipped, through physics and physiology and anatomy, to be able to act as if an optimal statistical decision-taker in the sense, for example, of de Groot (1970). In other words, an organism behaves as if a rational actor, with a specified utility function. In biology, it is common to refer to 'as-if optimality', to emphasise that consciousness is not implied by discussions of optimality. On the other hand, neither is it excluded—it is merely irrelevant. The construction of oak trees and salamanders *as if* they are rational actors raises the questions of whether humans have progressed to being *actually* rational actors, what that distinction might mean, and how one could tell. This, I take it, is the reason behind the very welcome invitation to write this target article for 'Biology and Philosophy'.

Acknowledgments I am very grateful to the editors and to all those colleagues who will write commentaries on this article. I am also grateful to Richard Gratwick and Paul Crewe for robust and relevant discussions over the past year, and to the Research Centre of St John's College, Oxford, for a grant (to AG and to Prof. Charles Batty) to pursue the project. Paul Crewe, Cécile Fabre, Andy Gardner, Richard Gratwick, Jarrod Hadfield, Tom Kemp, Samir Okasha, Stu West and Greg Wyatt commented very helpfully on an earlier draft.

References

Barrett R, Rogers SM, Schluter D (2008) Natural selection on a major armor gene in threespine stickleback. Science 322:255–257

Batty CJK, Crewe P, Grafen A, Gratwick R (in press). An individual fitness maximization principle in a class-structured population under uncertainty. J Math Biol
Cosmides LM, Tooby J (1981) Cytoplasmic inheritance and intragenomic conflict. J Theor Biol 89:83–129
Darwin CR (1859) The Origin of Species. John Murray, London
Davies NB, Krebs JR, West SA (2012) An introduction to behavioural ecology. Wiley-Blackwell, London
de Groot MH (1970) Optimal statistical decisions. McGraw Hill, New York
Dempster E (1955) The maintenance of genetic heterogeneity. Quant Biol 20:25–32
Edwards AWF (1994) The fundamental theorem of natural selection. Biol Rev 69:443–474
Ellner S, Rees R (2006) Integral projection models for species with complex demography. Am Nat 167:410–428
Engen S, Lande R, Saether B-E (2009) Reproductive value and fluctuating selection in an age-structured population. Genetics 183:629–637
Engen S, Lande R, Saether B-E, Dobson FS (1981) Reproductive value and the stochastic demography of age-structured populations. Am Nat 174:795–804
Ewens WJ (1989) An interpretation and proof of the fundamental theorem of natural selection. Theor Popul Biol 36:167–180
Ewens WJ (2004) Mathematical population genetics I. Theoretical introduction. Springer, Berlin
Fisher RA (1918) The correlation between relatives on the supposition of mendelian inheritance. Trans R Soc Edinburgh 52:399–433
Fisher RA (1930) The genetical theory of natural selection. Oxford University Press. See Fisher (1999) for a version in print
Fisher RA (1999) The genetical theory of natural selection. Oxford University Press. A Variorum edition of the 1930 and 1958 editions, edited by JH Bennett
Fisher RA (1958) The genetical theory of natural selection. 2nd ed. Dover Press, Mineola
Frank SA (1995) George Price's contributions to evolutionary genetics. J Theor Biol 175:373–388
Frank SA (1998) The foundations of social evolution. Princeton University Press, Princeton
Frank SA, Slatkin M (1990) Evolution in a variable environment. Am Nat 136:244–260
Gardner A, Grafen A (2009) Capturing the superorganism: a formal theory of group adaptation. J Evol Biol 22:659–671
Gardner A, West SA, Barton NH (2007) The relation between multilocus population genetics and social evolution theory. Am Nat 169:207–226
Grafen A (1984) Natural selection, kin selection and group selection. In: Krebs JR, Davies NB (eds) Behavioural ecology, 2nd ed. Blackwell Scientific Publications, Oxford, UK, pp. 62–84
Grafen A (1999) Formal Darwinism, the individual-as-maximising-agent analogy, and bet-hedging. Proc R Soc Series B 266:799–803
Grafen A (2000) Developments of Price's Equation and natural selection under uncertainty. Proc R Soc Ser B 267:1223–1227
Grafen A (2002) A first formal link between the Price equation and an optimisation program. J Theoret Biol 217:75–91
Grafen A (2003) Fisher the evolutionary biologist. J R Stat So Series D (The Statistican) 52:319–329
Grafen A (2006a) Optimisation of inclusive fitness. J Theor Biol 238:541–563
Grafen A (2006b) A theory of Fisher's reproductive value. J Math Biol 53:15–60 doi:10.1007/s00285-006-0376-4
Grafen A (2007a) Detecting kin selection at work using inclusive fitness. Proc R Soc Series B 274:713–719
Grafen A (2007b). The formal Darwinism project: a mid-term report. J Evol Biol 20:1243–1254
Grafen A (2007c) Inclusive fitness on a cyclical network. J Evol Biol 20:2278–2283
Grafen A (2008) The simplest formal argument for fitness optimization. J Genet 87:421–433
Hamilton WD (1964) The genetical evolution of social behaviour. J Theor Biol 7:1–52
Houston AI, McNamara J (1999) Models of Adaptive Behaviour: An Approach Based on State. Cambridge University Press, Cambridge
Kennedy JS (1992) The New Anthropomorphism. Cambridge University Press, Cambridge
Killingback T, Bieri J, Flatt T (2006) Evolution in group-structured populations can resolve the tragedy of the commons. Proc R Soc Series B 273:1477–1481
Kirkpatrick M, Johnson T, Barton N (2002) General models of multilocus evolution. Genetics 161:1727–1750

Lande R (2007) Expected relative fitness and the adaptive topography of fluctuating selection. Evolution 61:1835–1846
Lande R (2008) Adaptive topography of fluctuating selection in a Mendelian population. J Evol Biol 21:1096–1105
Lloyd D (1977) Genetic and phenotypic models of natural selection. J Theor Biol 69:543–560
Manolio T, Collins FS, Cox NJ, Goldstein DB, Hindorff LA, Hunter DJ, McCarthy MI, Ramos EM, Cardon LR, Chakravarti A, Cho JH, Guttmacher AE, Kong A, Kruglyak L, Mardis E, Rotimi CN, Slatkin M, Valle D, Whittemore AS, Boehnke M, Clark AG, Eichler EE, Gibson G, Haines JL, Mackay TFC, McCarroll SA, Visscher PM (2009) Finding the missing heritability of complex diseases. Nature 461:747–753
Maynard Smith J (1978) Optimization theory in evolution. Ann Rev Ecol System 9:31–56
Maynard Smith J, Price GR (1973) The logic of animal conflict. Nature 246:15–18
McGregor PK, Krebs JR, Perrins CM (1981) Song repertoires and lifetime reproductive success in the great tit (Parus major). Am Nat 118:149–159
Meszéna G, Kisdi E, Dieckmann U, Geritz S, Metz J (2001) Evolutionary optimisation models and matrix games in the unified perspective of adaptive dynamics. Selection 2:193–210
Moran PAP (1964) On the nonexistence of adaptive topographies. Ann Hum Genet 27:383–393
Ohtsuki H, Hauert C, Lieberman E, Nowak M (2006) A simple rule for the evolution of cooperation on graphs and social networks. Nature 441:502–505
Postma E, Spyrou N, Rollins L, Brooks R (2011) Sex-dependent selection differentially shapes genetic variation on and off the guppy Y chromosome. Evolution 65:2145–2156
Price GR (1970) Selection and covariance. Nature 227:520–521
Price GR (1972) Extension of covariance selection mathematics. Ann Hum Genet 35:485–490
Ridley M (1983) The explanation of organic diversity: the comparative method and adaptations for mating. Clarendon Press, Oxford
Rousset F (2004) Genetic structure and selection in subdivided populations. Princeton University Press, Princeton, New Jersey
Segestråle U (2000) Defenders of the truth. Oxford University Press, New York
Taylor PD (1990) Allele-frequency change in a class-structured population. Am Nat 135:95–106
Taylor PD (1996) Inclusive fitness arguments in genetic models of behaviour. J Math Biol 34:654–674
Taylor PD, Frank SA (1996) How to make a kin selection model. J Theoret Biol 180:27–37
Tinbergen N (1963) On aims and methods of ethology. Zeitschrift für Tierpsychologie 20, 410–433. This journal was renamed Ethology in 1986. The paper has been reprinted as Tinbergen (1985)
Tinbergen N (1985) On aims and methods of ethology. Animal Biology 55, 297–321. This journal was renamed Ethology in 1986. The paper is a facsimile of Tinbergen (1963)
van Doorn G (2009) Intralocus sexual conflict. Ann N Y Acad Sci 1168:52–71
West S, Griffin A, Gardner A, Diggle S (2006) Social evolution theory for microbes. Nat Rev Microbiol 4:597–607
West SA (2009) Sex allocation. Princeton University Press, Princeton
Williams GC (1966) Adaptation and natural selection. Princeton University Press, Princeton
Wright S (1969–1978) Evolution and the genetics of populations (Four Volumes). University of Chicago Press, Chicago