

The strategic gene

David Haig

Received: 21 November 2011 / Accepted: 10 March 2012 / Published online: 30 March 2012
© Springer Science+Business Media B.V. 2012

Abstract Gene-selectionists define fundamental terms in non-standard ways. Genes are determinants of difference. Phenotypes are defined as a gene's effects relative to some alternative whereas the environment is defined as all parts of the world that are shared by the alternatives being compared. Environments choose among phenotypes and thereby choose among genes. By this process, successful gene sequences become stores of information about what works in the environment. The strategic gene is defined as a set of gene tokens that combines 'actor' tokens responsible for an effect with 'recipient' tokens whose replication is thereby enhanced. This set of tokens can extend across the boundaries of individual organisms, or other levels of selection, as these are traditionally defined.

Keywords Gene selectionism · Environment · Phenotype · Developmental systems theory · Multi-level selection · Strategic gene · Relatedness

Gene selectionism is the conceptual framework that views genes as the ultimate beneficiaries of adaptations and organisms or groups as the means for genes' ends. Rival conceptual frameworks exist. Multi-level selection theory views genes as the lowest level of a nested hierarchy in which each level is subject to selection and each level can be the beneficiary of adaptations (Sober and Wilson 1994; Wilson and Sober 1994). Developmental systems theory similarly denies a privileged role for genes in development and evolution. In this framework, many things other than genes are inherited and many things other than genes have a causal role in development. It is the entire developmental system, including developmental

D. Haig (✉)

Department of Organismic and Evolutionary Biology, Harvard University, 26 Oxford Street,
Cambridge, MA 02138, USA
e-mail: dhaig@oeb.harvard.edu

resources of the environment, that reconstructs itself from generation to generation (Gray 1992; Sterelny and Griffiths 1999; Oyama 2000). Arguments among the proponents of these frameworks can be heated. Some of this argument is substantive but much is semantic. Different frameworks define fundamental terms in different ways. Without close attention to these semantic differences, substantive issues can become obscured by mutual incomprehension.

This article will explicate the non-standard definitions of gene, phenotype, and environment used by one particular gene-selectionist (the author). Although I strive for precision in definition, no definition will be unassailable. Natural selection constantly undermines rigid definitions because it is a process by which things of one kind become things of a different kind. Definitions themselves evolve and words acquire divergent meanings in different contexts. No language is unobjectionable if a hostile reader can choose how terms are interpreted. Mutual understanding should be facilitated by a clarification of how terms are used, but uniformity of definition is neither achievable nor desirable.

My intent is to clarify how central concepts should be defined to achieve a consistent gene selectionism, not to argue that gene selectionism is superior to other frameworks. Although I am a gene-selectionist by predilection, I have respect for, and sometimes use, multi-level selection theory. It is consistent and coherent. I also respect developmental systems theory, especially in its account of development. I see these frameworks as heuristic devices for thinking about evolutionary questions and alternative frameworks may be better suited to particular temperaments and particular questions.

Phenotypes

All that we mean when we speak of a gene for pink eyes is, a gene which differentiates a pink eyed fly from a normal one—not a gene which produces pink eyes per se, for the character pink eyes is dependent on the action of many other genes. (Sturtevant 1915)

Phenotypes have traditionally been defined as properties of organisms and thus must be redefined if phenotypes are to be considered properties of genes. *A gene's effects are its phenotype* (Dawkins 1982, p. 4). In this definition, an effect is simply a difference from what would be observed in the absence of the gene or in the presence of a variant gene, other things being equal. A gene considered in isolation does not have a phenotype. All assignments of phenotypes are based upon comparisons (Bouchard and Rosenberg 2004). These comparisons are implicit in the measurement of genic fitness, a phenotype of particular importance, as a change in relative allele frequency. Thus, a gene's phenotype depends on the implicit or explicit alternative with which it is compared.¹ For some evolutionary questions, comparisons are made between existing variants of a known sequence. For other

¹ A gene's effects are often measured relative to the average for all alternatives (alleles) in the population. Fisher (1941) defined the average effect as the partial regression of the phenotypic measure on the presence or absence of the gene.

questions, an existing gene might be compared to a hypothetical alternative, or a hypothetical gene and hypothetical alternative might be posited as making a specified difference in the world without tying the difference to a particular genetic locus.

Consider a sequence variant that has become fixed in a population by a selective sweep. The variant allele was initially present in a single copy in a single cell. Early in the sweep, the allele's increase in frequency was determined by its effects relative to the previously established allele. Once the sweep is complete, the allele maintains its high frequency, or is itself replaced, depending on its effects relative to new mutations. Some mutations will result in complete loss of function. In this case, the phenotype that is subject to selection is the difference between a functional and nonfunctional allele. Other mutations will cause expression in a new cell type, a change in alternative splicing, a change in promoter activity, etc. For each kind of mutation, there is a different phenotype under selection.

The definition of phenotype as a gene's effects changes how environmental and genetic factors are conceptualized. Phenotype is no longer seen as a sum of genetic and environmental influences (plus interaction terms). All phenotypes are effects of genes but a gene's phenotype may encompass different effects in different environments (the gene's reaction norm) and includes the gene's effects on the environment. If there are differences among organisms that are uninfluenced by genes, then these differences are part of no gene's phenotype. If a frog has a leg amputated by some purely random event, then the absence of the limb is not a phenotype, although how the frog copes, or fails to cope, with the amputation may be part of the phenotype of many genes. We are all interactionists. A definition of phenotype in which all effects are ascribed to genes, but genes' effects may vary among environments, has much to recommend it if it helps us move beyond sterile debates about nature versus nurture.

A gene achieves its effects by interacting with proteins, RNAs, DNA, and other molecules. The other molecules with which it interacts can be conceptualized as part of the gene's environment. Of particular importance are a gene's interactions with RNA polymerases in the process of transcription. However, a gene's phenotype need not be mediated solely by the coding and non-coding RNAs transcribed from its sequence. DNA sequences can adopt multiple conformations depending on conditions in the nuclear environment and these conformations can influence whether the gene is transcribed. For example, the 5'-flanking region of human γ -globin genes responds to low pH by forming an intramolecular triplex in which a purine-rich strand inserts into the major groove of the contiguous double-helix leaving an unpaired pyrimidine-rich strand. Point mutations that destabilize this structure are associated with hereditary persistence of fetal hemoglobin (Bacolla et al. 1995). Thus, the pH-dependent ability to fold back upon itself and thereby inhibit its own transcription can be considered part of the phenotype of the γ -globin gene. As another example, the imprinted *H19* gene interacts directly (without RNA intermediates) with imprinted regions on other chromosomes and influences when in the cell cycle these other genes are replicated (Sandhu et al. 2009).

Genes have been viewed as catalysts that facilitate chemical reactions but are not changed by those reactions (Haig 1997). Much recent attention however has focused

on chemical reactions that alter the gene itself. For example, a DNA sequence may be ‘epigenetically’ modified by its interaction with methyltransferases. If alternative DNA sequences exist that are not subject to methylation, then methylation is part of a gene’s phenotype compared to those alternatives. If a methyl group, once attached, is faithfully inherited and more or less permanent, then a methylated and unmethylated sequence can be considered alternative genes compared to each other. But, if a genetic lineage switches back and forth between methylated and unmethylated states, then the alternative states are better considered a reaction norm and part of the gene’s phenotype compared to alternative alleles that do not undergo switching (Haig 2007).

Functions and side-effects

A gene may have effects which influence the probability that it will be replicated. Genes that promote their own replication will be perpetuated, whereas alternative genes that are less effective replicators will be eliminated. The *effects* of a DNA sequence may thus be included among the *causal* factors that account for the presence of the sequence in a gene pool. It is this causal feedback between genotype and phenotype—when combined with a source of genetic novelty (mutation)—that explains how a purposeless process (natural selection) can produce purposeful structures and functions (adaptation). The environment selects among phenotypes *and thereby selects among genes*. By this means, gene sequences come to embody and represent ‘information’ about what works in the environment (Shea 2007; Frank 2009).

The effects of a gene can be classified as either *functions* (effects that are beneficial for the gene) or *side-effects* (effects that are neutral or harmful for the gene). A gene’s *functions* consist of those of its effects that have contributed, however indirectly, to its own transmission from past generations. In so far as the future repeats the past, such functions will contribute to the gene’s transmission to future generations. All effects of a gene comprise its phenotype and are subject to selection, but only those effects that promote a gene’s replication comprise its functions (Haig and Trivers 1995, p. 27). There is selection *for* a gene’s functions but selection *of* its side-effects.² If there is to be selection of a harmful effect then it must be associated with an even greater beneficial effect (a selective trade-off). Effects are judged as functions or side-effects by their average contribution to replication over many occurrences not from a single occurrence.

Functions are the adaptations of genes. For an effect to qualify as a function, variant genes must have been eliminated in the past because they lacked the effect and, if the effect is to remain a function, such variants must continue to be eliminated when they arise. A gene’s effects can change when the environment changes and a given effect can shift its status from function to side-effect or the

² A function is synonymous with a selectional property of a gene (Shea 2007, p. 320) or the property *for* which there is selection (Sober 1984, pp. 99–100). In Sober’s example of small green balls passing through a sieve, there was selection *for* the property of smallness and selection *of* objects that were small and green. He did not state whether there was selection *of* the property of greenness. I consider there to be selection *for* and *of* functions of a favored gene but simply selection *of* side-effects.

reverse. There is no inconsistency in saying that an effect was once a function but is no longer. Teleological language is appropriate when referring to the functions of genes because functions are final causes. They are both causes of a gene's persistence and effects of the gene.

Environments

A gene's environment encompasses all factors that are shared with the alternative against which the gene's effects are measured. It contains not only factors external to the cells and bodies of organisms, but also (and more immediately) these cells and bodies themselves. A body can be viewed as the collectively-constructed niche of the genes of which it is an extended phenotype. Among the most important parts of a gene's environment are the other molecules with which it interacts. Other genes, even other alleles at the same locus, are parts of a gene's social environment (Fisher 1941; Sterelny and Kitcher 1988; Okasha 2008). On the other hand, any factor that is experienced by a gene, but not by its alternative, belongs to the gene's phenotype, not its environment. Phenotypes and environments can be defined similarly for other determinants of difference but it is genes that are the focus of this paper.

Genes may have effects that vary in different environments, in adaptive or non-adaptive ways. If alternative genes experience a similar range of environments but exhibit different responses to these environments, then their reaction norms are phenotypes subject to selection. Genes may also have effects that modify the environment or phenotype of other genes.

Genetic inheritance and phenotypic development can be conceptualized as orthogonal axes (Bergstrom and Rosvall 2011). The vertical axis represents transmission of genetic information from progenitors to progeny whereas the horizontal axis represents processes of development within generations. Relations between genes and environment differ on the two axes. The environment has primacy on the vertical axis. All information in genes was put there by the environment via selection among alternative genetic effects. But, on the horizontal axis, genes and environment interact to create form. Neither has explanatory primacy. On the vertical axis, genes refer to *past* environments whereas, on the horizontal axis, genes interact with the *current* environment.

Genealogical (vertical) and ontogenetic (horizontal) axes are not causally isolated. If a gene makes a selective difference in the current environment, this will be reflected in a changed composition of the gene pool. In any particular generation, entire genomes are selected and one cannot ascribe a selective difference to a particular gene. However, genomes are disassembled and reassembled in each generation by processes of recombination. Therefore, over a series of generations, shorter sequences of DNA are tested against multiple genetic backgrounds. On the ontogenetic axis, the effects of genes are highly non-additive, because of complex interactions with the environment (which includes other genes), but, on the genealogical axis, sustained changes in gene frequency are explained by the average additive effects of smaller parts of the genome (Fisher 1941; Ewens 2011).

Are genes dispensable?

To recapitulate, a gene's world can be divided into phenotype and environment. The phenotype contains those parts of the gene's world that are different from the world of the alternative to which it is compared whereas the environment contains those parts of the gene's world that are shared with the alternative's world. The environment chooses among the phenotypes of alternative genes and thereby chooses among genes. Could a simpler story be told in which the environment selects among heritable phenotypic differences without the need to invoke heritable determinants of difference?

The principal reason for invoking determinants is that causality matters. Haig (2003; see Godfrey-Smith 2009, p. 171) considered a model in which heritability of birth weight was high and different birth weights were associated with different probabilities of survival, but there were no fitness differences associated with genetic causes of variation. Rather, fitness differences were associated solely with environmental contributions to the variance. Perhaps mothers of all genotypes produce heavier and healthier babies in better environments. If so, there would be the appearance of directional selection for heavier babies but no response to selection. Perhaps greater environmental perturbations from genotype-specific optimal birth weights result in reduced survival of babies. If so, there would be the appearance of stabilizing selection but no reduction in genetic variance. In both scenarios, birth weight and fitness are correlated and birth weight is heritable but fitness is not. Heritability of a trait that is correlated with fitness is not enough. Natural selection requires heritability of the *causes* of variation in fitness.

The second reason for invoking genes, as heritable determinants of difference, is that genes can be considered agents that benefit from the phenotypes they cause. This agential view will be developed below with the concept of the gene as strategist.

What is the gene-selectionist's gene?

The discussion so far has put off consideration of how genes should be defined to qualify as the beneficiaries of adaptations. Two kinds of questions will be addressed. The next section addresses the spatial extent (measured in base pairs) of the evolutionary gene. A subsequent section considers the distinction between gene types and gene tokens and defines the strategic gene as a set of tokens of an evolutionary gene.

Evolutionary gene concept

Williams (1966, p. 24) and Dawkins (1976, p. 30) defined a gene as a rarely recombining stretch of DNA that is transmitted intact over multiple generations. Evolutionary genes include stretches of DNA that are 'context-sensitive difference makers' (Sterelny and Griffiths 1999, p. 87) but need not be restricted to that category. The existence of DNA sequences without effects has been presented as a problem for the evolutionary gene concept (*ibid*; pp. 79, 85, 92) but the difficulty evaporates if it is conceded that an evolutionary gene need not be subject to

selection. Nor is there a problem if different DNA sequences have identical effects (ibid; p. 89). One can still identify them as different genes on the basis of their sequence. A gene without effects is a gene without a phenotype. Such genes make no selective difference although their relative frequency may change by genetic drift or draft. Similarly, a gene that has erratic effects on its own replication, such that there is no average effect, is not subject to selection, but is still a gene.

The linear extent of an evolutionary gene can be considered to be defined by the distance along a chromosome over which there is strong linkage disequilibrium. Consider two neighboring DNA segments X and Y. If $P(X)$ and $P(Y)$ are the relative frequencies of X and Y, and $P(XY)$ is the frequency with which X occurs together with Y, then the distributions of X and Y are statistically non-independent if $P(XY) \neq P(X)P(Y)$. In this case, knowing whether X is present provides information about Y, and knowing whether Y is present provides information about X. Such non-independence of DNA segments is known as linkage disequilibrium. In some cases, blocks of high linkage disequilibrium may be separated from each other by breaks at 'hot spots' of recombination. But in many cases, linkage disequilibrium declines gradually with distance so that a chromosome cannot be divided into discrete evolutionary genes. Rather, there is a region of high linkage disequilibrium associated with each polymorphic site that can be considered the evolutionary gene with respect to that site, but evolutionary genes defined with respect to different polymorphic sites may overlap.

If X and Y are perfectly associated, then natural selection is indifferent to whether an effect (relative to non-X and non-Y) is due to X or Y alone, to the sum of their effects, or to their interaction. (The causal question can be addressed by experimentally breaking the association and producing X without Y or Y without X.) By contrast, if X and Y are randomly associated, then X can be treated as a variable part of the environment of Y and Y as a variable part of the environment of X. Clearly, the extremes of perfect and random association are the ends of a continuum. As linkage disequilibrium increases, it becomes more convenient to treat X and Y as parts of the same evolutionary gene. As linkage disequilibrium decreases, it becomes more convenient to treat X and Y as part of each other's environment.

Godfrey-Smith (2009, pp. 135–139) believes the inability to assign non-arbitrary boundaries to evolutionary genes is a major flaw of gene-selectionism. In his view, "A Darwinian population is made up of a set of definite countable things" but evolutionary genes are, at best, marginal Darwinian individuals because they fail to satisfy this 'definite countable' criterion. He concedes that the lack of countability does not matter much "if one's point of view is sufficiently pragmatic" and that similar problems sometimes arise in counting organisms, cells, and groups, but the difficulties are, in his view, particularly pronounced for evolutionary genes. They are not the sort of "real entities that undergo the kind of change that Darwin described." Rather than talk of genes, "In an evolutionary context it is more accurate to talk of *genetic material*, which comes in smaller and larger chunks, all of which may be passed on and which have various causal roles."³

³ Many significant things do not have precise boundaries. No line on the ground demarcates where the Rocky Mountains begin (nor is there any way to count the number of peaks in Colorado without making

Godfrey-Smith confounds two kinds of count in his calculation of the number of genes in a bacterial population. First, he counts the number of genes in a bacterium (a few thousand). Second, he counts the number of bacteria in the population (a million). Then, he multiplies these numbers to obtain the number of genes in the population (a few billion). The first number is a count of different kinds of genes. This number is poorly defined because boundaries between genes are indeterminate, but the items so counted do not constitute a Darwinian population. The second number is a count of how many genes of each kind. This is the size of a Darwinian population, but the count is not affected by where one places the boundaries between genes.

Similar issues arise when counting evolutionary genes in sexual eukaryotes. The number of genes on a chromosome is poorly-defined because boundaries between genes are fuzzy, but the number of copies of the X chromosome in a group of organisms, or the number of 'genes' at a particular *site* on that chromosome (the counts should be the same), is not affected by where one places boundaries between genes. Only the latter number measures the size of a Darwinian population. If the population at a particular site consists of different variants ('alleles'), then selection can be measured by changes in the non-arbitrary numbers of these variants. Linkage disequilibrium is a measure of how representative the count at one polymorphic site is of counts at nearby polymorphic sites.

Evolutionary genes are conceptually central for Williams and Dawkins because such stretches of DNA persist over many generations whereas organisms, cells, and groups are ephemeral. Within their conceptual framework, discrete boundaries are unimportant but persistence is central. Within Godfrey-Smith's conceptual framework, Darwinian individuals should be clearly identifiable things but persistence is not one of their essential properties. In the interests of pluralism, I ask that judgment on these issues be temporarily put to one side in the interest of explicating gene-selectionism in its own terms. I expect my definitions of fundamental terms will be contested by critics of gene-selectionism once these definitions are clearly presented.

Expanding the evolutionary gene

DNA blocks defined by strong linkage disequilibrium need not respect the boundaries of protein-coding units. They may be smaller or bigger than such units. As George Williams (1966, p. 24) noted, "Various kinds of suppression of recombination may cause a major chromosomal segment or even a whole chromosome to be transmitted entire for many generations in certain lines of descent. In such cases the segment or chromosome behaves in a way that approximates the population genetics of a single gene." From this perspective, a

Footnote 3 continued

pragmatic, somewhat arbitrary, choices about what counts as a peak). One could, if one chose, think of North America as comprised of smaller and larger chunks of *landscape material* without naming any topographical features with fuzzy boundaries. Locations and areas could be identified by latitude and longitude, but this would be cumbersome. Natural selection is pragmatic rather than principled (whatever 'works' works) and my own approach to evolutionary theory is similarly pragmatic.

mitochondrial genome or the non-recombining portion of the Y chromosome can be considered a single evolutionary gene, as can the genome of an asexual organism.

The definition of linkage disequilibrium, $P(XY) \neq P(X)P(Y)$, can be generalized to all cases of non-independence between X and Y. From this perspective, species boundaries are a major cause of linkage disequilibrium. For example, there is complete linkage disequilibrium between grey squirrel and red squirrel DNA (although polymorphic sites may be randomly associated within each species). The DNA of the introduced grey squirrel has rapidly displaced the DNA of the indigenous red squirrel from most British forests with the frequency of all parts of the red squirrel genome changing in concert relative to all parts of the grey squirrel genome. Some DNA segments from red squirrels might be selectively favored in grey squirrel bodies, but such combinations never arise. Natural selection acts on phenotypic differences between the gene pools but does not 'see' independent effects of smaller DNA segments. Ecological displacement can be considered a selective process in which the non-recombining units are the gene pools of the competing species (Williams 1986; Haig 1997). Proponents of the evolutionary gene concept could accommodate interspecific competition by identifying the gene pools as 'evolutionary genes' or could avoid the issue by restricting the concept's application to natural selection within sexually-recombining populations.

Types, tokens, and strategic genes

What is the selfish gene? It is not just one single physical bit of DNA ...it is all replicas of a particular bit of DNA distributed throughout the world. (Dawkins 1976, p. 95)

A proper understanding of the units of selection problem must take account of an important symmetry: *Just as organisms are parts of groups, so genes are parts of organisms.* (Sober and Wilson 1994, p. 549)

A gene that is distributed throughout the world cannot be part of an organism that is localized in space. Whether genes are the 'unit of selection' has remained contentious, in part, because different meanings of 'gene' are conflated (Haig 1997). A first step to untangle this knot is to recognize that 'gene' can refer both to a type and to tokens of the type (as well as to collections of tokens of a type). Gene tokens are physical objects but gene types are abstract kinds. It is tempting to simplify matters by suggesting that Dawkins refers to the type when he describes a gene as all replicas of a particular bit of DNA whereas Sober & Wilson refer to tokens when they identify genes as parts of organisms. However, such an attempt to cut the knot fails for the 'selfish gene' because universal benevolence is not predicted once a gene has gone to fixation (Haig 2006).

Evolution is often characterized as changes in gene frequency and the phenotypic effects of these changes. Changes in frequency imply counting but gene tokens are rarely counted, rather population geneticists usually lump together large numbers of tokens defined by the boundaries of individual organisms and count these *sets* as a single gene (Queller 2011). Thus, all tokens of a type in a haploid individual are counted as one gene (haploids have one allele) whereas all egg-derived tokens in a

diploid individual are counted as one gene and all sperm-derived tokens as another gene (diploids have two alleles). This sleight-of-hand facilitates the development of simple mathematical models of evolutionary change.

Multi-level selection theory implicitly defines a gene as a single token within a cell in some contexts but as the set of all tokens of a type within an organism in other contexts. Gene-selectionists also implicitly define the gene as a set of tokens, but a set that may be distributed across multiple organisms, for example across multiple members of a hive. Haig (1997, 2006) has called this set of tokens the *strategic gene* because it is the unit that can be considered a strategist in an evolutionary game played with other strategic genes.

A gene token is transcribed when an RNA polymerase copies its sequence into a functional RNA (such as a messenger RNA that is translated into a protein) and is replicated when a DNA polymerase takes apart the two strands of its double helix and uses each as a template to produce two new tokens. The strategic gene groups together tokens that cause an effect (actors) with tokens of the same type whose probability of replication is thereby affected (recipients). Actors may be located in somatic cells of a multicellular organism with recipients located in germ cells of the same organism, but actors and recipients may also be tokens of the same type located in different organisms. The strategic gene is not a fixed entity but can evolve to encompass more, or fewer, tokens of its type.

Consider a particular token in a germ cell (the focal token) and trace its ancestry back to the *urtoken*, the very first token of its type to arise by mutation. From the urtoken, a dichotomously-branching tree can be envisaged that represents the history of all tokens of the type, with the focal token at one of the tips of the tree.⁴ The path through this tree from focal token to urtoken summarizes the selective history of the focal token (Fig. 1a). In organisms with a strong germ-soma distinction, the tokens on this path are located in germ cells whereas most of the tokens on lateral branches are located in somatic cells. Tokens on the germ-path may be recipients of effects from tokens on lateral tips of the token-tree (Fig. 1b, c). Selection acts when these effects make a replicative difference (cause a change in relative frequency) relative to tokens of some other type. In this schema, phenotypic effects flow 'inward' from somatic actors to germ-line recipients. These 'causal arrows' influence which tokens are replicated but do not change the type of the token.

The extent of the strategic gene is determined by the number of replication cycles that separate the tokens responsible for a phenotypic effect from the tokens that thereby gain a selective advantage. Tokens on remote tips of a token-tree may be unable to exert selective effects on each other's replication because token-trees are broken up into selectively-isolated fragments by spatial dispersal of tokens and mixing with tokens of other types. Selectively-isolated tips of a token-tree belong to different strategic genes. Which tokens belong to a strategic gene is determined by the answer to the question, what is it about the effects of tokens of this type that accounts for a focal token being present in the population rather than a token of

⁴ Dichotomous branching is a property of the semi-conservative replication of DNA but is not a necessary feature of the genealogies of genes. Rabies virus has a single-stranded RNA genome. This strand is transcribed to produce the complementary single-stranded antigenome that serves as a template to transcribe multiple copies of the parent genome (Wunner 2007).

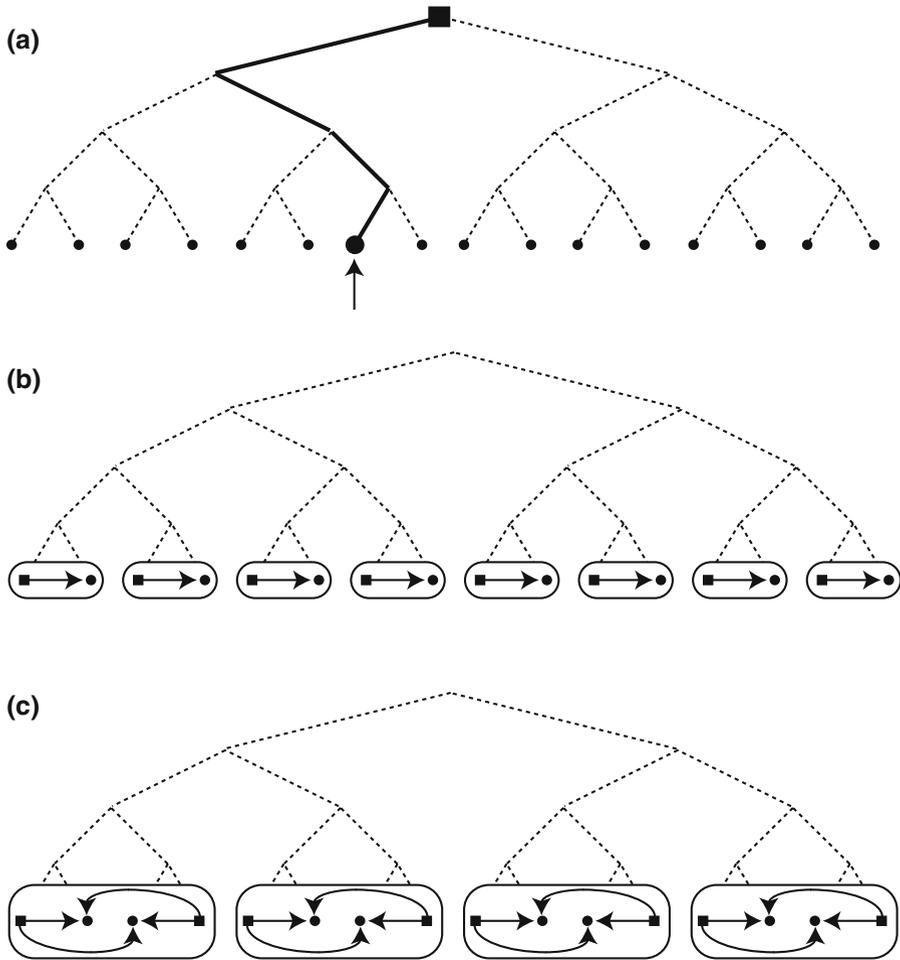


Fig. 1 Representations of a token tree: **a** Filled circles represent tokens of a gene type. The arrow identifies the focal token. Its ancestry can be traced back to the first token of its type (the urtoken represented by a square). The token-tree represents the genealogical relationships of all tokens descended from the urtoken. **b** A simple example in which an actor-token (square) confers a benefit (arrow) on a recipient-token (circle) of the same type. Groups of tokens that interact in this way correspond to strategic genes (enclosed in cartouche). **c** The strategic gene can evolve to encompass more (or fewer) tokens of a type

another type. These ideas are explored below with four illustrative cases chosen to mimic a progression of multi-level selection theory from selection among cells (Case 1), to selection among multicellular organisms (Case 2) to selection among groups of organisms (Cases 3 and 4).

Case 1: Planktonic alga

Consider a haploid alga in which mitosis directly follows DNA synthesis. Daughter cells separate and rapidly become mixed with unrelated cells. Each cell contains a

single token of each evolutionary gene. After DNA replication, the cell contains two tokens that immediately part company and are mixed with tokens from the general population by turbulent flow. It is as if the tips of the token-tree break free after each round of DNA replication and are randomly mixed with the tips of other token-trees. If two cells subsequently met, and a token in one acted to enhance the replication of the token in the other, then the beneficiary would be a random member of the pool of tokens from the perspective of the actor. Thus, the action would be without selective consequence apart from any correlated effects on the actor's own replication. The strategic gene corresponds to a single token in a single cell. This token is both an actor and the recipient of its own effects.

Case 2: Fish with planktonic larva

Now consider a fish species in which diploid larvae are mixed by turbulent flow after eggs are fertilized. Each zygote contains two unrelated tokens (random samples from the token pool). These tokens replicate in synchrony to produce a fish body containing many tokens of the two parental types. All the descendants of a zygotic token can be considered a strategic gene, with the descendants of the other zygotic token part of that gene's environment. Thus, the fish's body contains two strategic genes at each locus. These may be fragments of different token-trees or disjunct parts of the same token-tree (descended from the same urtoken). Actor and recipient roles of each strategic gene are divided between tokens in somatic cells (actors) and tokens of the same type in germ cells (recipients).

The strategic gene does not extend beyond an individual fish because encounters between fish are with unrelated individuals (random samples from the fish pool). It is as if token-trees are broken into fragments (strategic genes) each of which has at its origin a single token in a fertilized egg and at its tips the myriad tokens of one of the parental alleles in a fish body. Actor-recipient interactions among tips of a fragment have selective consequences, but interactions among tips of different fragments do not, because the fragments of different token-trees are well-mixed.

Case 3: Huddling with strangers

Consider a group of unrelated mice huddled together on a cold night. Non-shivering thermogenesis in brown adipocytes consumes fat to generate heat thus increasing the chances that each member of the group survives the night. Strategic genes are confined to individual mice, with actors in each mouse's brown adipocytes and recipients in the same mouse's germ cells. Heat generation benefits tokens in the germ cells of other mice but this does not constitute a selectable difference because those tokens are a random sample from the token pool. Huddling behavior has been selected because actors in each mouse's central nervous system cause their mouse to seek a favorable environment next to another warm body. In this example, there is a hierarchy of groups of gene tokens: tokens are grouped into strategic genes; strategic genes are paired in diploid mice; and mice huddle in groups.

Case 4: Huddling with kin

Now suppose that the huddling group is an outbred litter of mouse pups. The egg-derived tokens of each pup trace their ancestry to one of the two tokens at a locus that were present in the zygote from which the litter's mother developed. For each of these ancestral tokens, the other was a random sample from the token pool. Thus, the litter can be considered to contain a group of two egg-derived strategic genes, each of which has actors in brown adipocytes and recipients in germ cells of multiple pups. Heat generation by actors in one pup benefit recipients in germ cells of other pups as well as in germ cells of their own pup.

By similar reasoning, a litter fathered by n unrelated males will contain up to $2n$ sperm-derived strategic genes. Therefore, litters contain more sperm-derived strategic genes than egg-derived strategic genes when $n > 1$. Sperm-derived strategic genes should favor less heat generation than egg-derived strategic genes because thermogenesis provides a public good (heat) with private costs (lipids burnt) and larger groups are less effective than smaller groups at providing public goods (Olson 1961). This prediction is supported by evidence that genes of paternal origin inhibit non-shivering thermogenesis in brown adipocytes (Haig 2008, 2010).

Although each strategic gene is distributed across multiple members of the litter there remains a special relationship between actors and recipients in the same pup. From the perspective of an actor in a particular pup, the germ cells of that pup definitely carry copies of the actor but the germ cells of littermates only have a chance of carrying its copies by recent common descent. Thus, effects on the replication of tokens of a strategic gene in germ cells of littermates are 'diluted' with effects on tokens of other strategic genes. The amount of heat produced is predicted to be less than in a group of clonal individuals but more than in a group of unrelated individuals.

Relatedness

The mixing of tokens of different types limits the reach of strategic gene action. Once fragments of different token-trees are randomly mixed then a strategic gene can be no larger than an individual fragment. In the first three cases, there was complete mixing of algal cells, fish larvae, or mice in a 'planktonic' phase. In the fourth case, there was partial mixing of fragments after meiosis. Every littermate contained a continuation of one of the two token-trees in its mother but, from the perspective of an actor in a particular mouse, it was impossible to tell which littermates received which kind of token. It is as if tokens in the actor's own mouse were full-strength but egg-derived tokens in litter mates were diluted by half with tokens randomly chosen from the pool of all tokens. This dilution factor can be conceptualized as the *relatedness* of tokens that received a cost or benefit to the tokens that conferred the cost or benefit. Such dilution of an actor's type among recipients of the actor's effects can be achieved in numerous ways, including dispersal of tokens away from their relatives, the uncertainty of meiotic segregation, and the mixing of sperm-derived tokens from different males among the progeny of polyandrous females.

Multi-level gene-selectionism

From a gene-selectionist perspective, the strategic gene of Case 1 corresponds to a single token that is both actor and recipient; the strategic genes of Case 2 and 3 consist of a large group of tokens within a single organism, with a separation of roles between somatic actors and germ-line recipients; and the strategic gene of Case 4 is distributed across a group of organisms with somatic actors and germ-line recipients in multiple member of the group. When viewed in this light, gene-selectionist and multi-level selection frameworks are fundamentally similar ways of describing the same set of phenomena.

The strategic gene navigates a perilous path between the Scylla of the gene token (material gene) and the Charybdis of the gene type (informational gene). It is a set of tokens but not the set of all tokens of a type. Its tokens may be distributed across multiple levels of the hierarchy of interactors of multi-level selection theory, but it is not a level of this hierarchy. The strategic gene combines tokens responsible for a phenotypic effect with tokens of the same type that benefit, directly or indirectly, from the effect.⁵ It is the beneficiary of the effects it causes. As such, it is a *unit of adaptive innovation* and a unit of self-interest.

Genetic tokenism

When the present paper was near completion, Gardner and Welch (2011) published *A formal theory of the selfish gene*. Here I will briefly identify some important differences between our approaches. G&W identify the selfish gene with the gene token rather than the gene type (allele). They considered, but rejected, a definition of genes as sets of tokens in the following terms: “There may be viable alternatives that recover the distributed agent view, such as seeing the gene as a cloud of identical-by-descent scraps of nucleic acid. However, such agents are statistical rather than concrete objects, and it is difficult to assign diffuse probabilistic clouds a causal role in evolutionary biology.”

In the formalism of G&W, the network of interactions among gene tokens is assumed to be isomorphic for every token. Each focal token is a recipient of effects from other tokens performing every role in its allele’s strategy and the focal token itself performs every role in its strategy with respect to some other token. The assumption of isomorphism has mathematical elegance but collapses everything a gene might do into a single token that can be considered the recipient of its own effects. By contrast, the strategic gene is a cloud of identical-by-descent tokens with the possibility of a division of labor among its parts. For G&W, the social gene is a token that may be selfish, altruistic, or spiteful with respect to other tokens, including identical-by-descent tokens of the same type, whereas, in my formalism, the actions of a strategic gene are always judged by the criterion of the collective self-interest of its tokens considered as a group.

⁵ A strategic gene could also consist of tokens (actors) that reduced the replication of tokens of the same type (recipients), but in this case the strategy is inept and the gene would be eliminated by natural selection.

Developmental systems framework

Developmental systems theory has been presented as a radical challenge to gene-centered accounts of development and, by extension, to gene-centered accounts of evolution (Gray 1992; Griffiths 1998; Oyama 2000). In the developmental systems framework, genes are just one among many components of a developmental matrix and have no privileged causal role in development. Gene selectionism is seen as misguided because it implicitly endorses a dualistic account of development in which genes are the carriers of preformed instructions of how to construct phenotypes, with the environment cast in a subsidiary and passive role.

Adaptation by natural selection take a back seat to ontogenetic questions for proponents of developmental systems theory whereas adaptation occupies the front seat for proponents of gene selectionism. A simple resolution of their disputes would be to propose that the two frameworks address different questions. Such a proposal would probably be perceived as partisan because the premise that ontogenetic and evolutionary questions require different kinds of answer is generally accepted by one side (Dawkins 1982, p. 98; Williams 1986) but rejected by the other (Gray 1992, p. 187; Oyama 2000, p. 45). Gene-selectionists believe that a conceptual separation of developmental from adaptive explanations aids clarity of thought whereas developmental-systemists believe such separation obscures more than it illuminates.

My intent has been a defense of the territory of gene-selectionism, not an attack on the core territory of developmental systems theory. Developmental-systemists are on firm ground when they argue that genes do not have a privileged role in a causal account of development, and that phenotypes (in the traditional sense) are constructed by complex, highly non-additive, interactions of genetic and environmental factors (Gray 1992; pp. 172–174). For these reasons, effects of individual genes on the course of development cannot be isolated, but this is exactly what natural selection does over the course of many generations as genes are tested in different genetic backgrounds and in a series of environments. Natural selection extracts the average additive *effects* of genes as the environment ‘chooses’ among phenotypes.

In the foreword to the second edition of Oyama’s *Ontogeny of information*, Lewontin (2000) remarked:

Throughout the history of modern biology there has been a confusion between two basic questions about organisms: the problem of the origin of differences and the problem of the origin of state. At first sight these seem to be the same question, and taken in the right direction, they are. After all, if we could explain why each particular organism has its particular form, then we would have explained, *pari passu*, the differences between them. But the reverse is not true. A sufficient explanation of why two things are different may leave out everything needed to explain their nature. (p. viii)

By implication, understanding causes of difference is subsidiary to understanding causes of state. Lewontin (1974, 2000) and Oyama (2000, pp. 52, 155) have

perceptively identified an issue on which there is conceptual disagreement between the rival camps.⁶

In the developmental systems framework, genes are non-privileged components of a developmental matrix and it is a conceptual error to assign phenotypic features to genes rather than to the matrix as a whole. The entire matrix, or life cycle, constructs itself epigenetically in each generation. A gene's environment, as I have defined it in this paper, is co-extensive with the developmental matrix except that the environment is defined to exclude the gene. There are effects of the gene and all else is environment.

Effects are differences. Natural selection chooses *this* versus *that* based on some difference. Why we observe these developmental systems and not others is explained, in part, by a long history of selection among differences and thereby selection of particular heritable difference-makers. The selection of differences can result in profound changes of state. One might say that there is selection *for* the causes of differences and selection *of* the causes of state.

Gene-selectionists use the language of statistics, of variances, correlations and average effects, whereas developmental-systemists prefer 'causal' accounts. The contrast, within physics, between mechanics and thermodynamics provides a useful analogy. Thermodynamics is a statistical theory, not an exact causal theory. It makes predictions that are right on average. In principle, a thermodynamic account of any system could always be superseded by a complete mechanical account, but in many circumstances an exact causal account is not practical, nor even possible, nor would it add much to the thermodynamic explanation.⁷

There is, in principle although not in practice, a complete account of all evolutionary change expressed in terms of proximate physical causes that makes no appeal to concepts of selection, information, average effects, and the like. But, I will settle for what is practical and predictive. Sober (1984, p. 311) writes "The strategy of averaging over contexts is the magic wand of genic selectionism. It is a universal tool, allowing *all* selection processes, regardless of their causal structure, to be represented at the level of the single gene." I agree, but see this as a strength rather than a weakness of gene selectionism.

A developmental system exists in which thornbill chicks in a nest are fed by a family of thornbills. This system is reconstructed in each generation as part of the life cycle of thornbills. The nest is a key developmental resource that is constructed afresh in each generation. There is another developmental system in which a cuckoo chick, in a similar nest, is fed by a family of thornbills until the cuckoo is larger than

⁶ Isadore Nabi (pers. comm.) offers the rejoinder "If we could explain how each organism has evolved its particular form, by the selection of differences, then we would have explained, *per stirpes*, why it has its particular state. But the reverse is not true. A sufficient explanation of how an organism develops may tell us nothing about why it has its particular form."

⁷ Fisher (1958, p. 39), a statistician *par excellence*, compared his fundamental theorem of natural selection to the second law of thermodynamics. Dynamics exhibits 'time-reversal invariance' whereas thermodynamics is not time-reversible. Processes of mutation, drift and recombination are time-reversible at the micro-level of the individual gene token but increase entropy at the population level. Fisher believed that natural selection was associated with a unidirectional arrow of time that tended to increase order. Okasha (2008, p. 346) rejects the existence of such an arrow because increase in 'adaptation' is countered by deterioration in the environment.

a thornbill. This system, including the nest, is reconstructed in each generation of the thornbill-cuckoo symbiosis. Developmental resources are similar in the two systems. The key difference-maker is the placement, in the nest, of a cuckoo egg with its cargo of cuckoo genes. Developmental-systemists see these systems as fundamentally similar whereas gene-selectionists see them as fundamentally different. Both are right. The origins of adaptation, by selection for the causes of average additive differences, and processes of development, involving highly non-additive interactions among causes of state, are both fundamental questions.

Are genes special?

Many things besides genes are replicated, including membranes, song traditions, burrows, and nests (Sterelny et al. 1996; Sterelny and Griffiths 1999, p. 70).⁸ Genes however possess a peculiar property that distinguishes them from most other heritable difference-makers. Muller (1934) observed that genes were autocatalysts:

But the most remarkable feature of the situation is not this oft-noted autocatalytic action in itself—it is the fact that, when the structure of the gene becomes changed, through some ‘chance variation,’ the catalytic property of the gene may become correspondingly changed, in such a way as to leave it still autocatalytic. In other words, the change in gene structure—accidental though it was—has somehow resulted in a change of exactly appropriate nature in the catalytic reactions, so that the new reactions are now accurately adapted to produce more material just like that in the changed gene itself. (page 34)

This prescient passage was written before elucidation of the structure of DNA. We now understand, in considerable molecular detail, how this extraordinary property is achieved by the two strands of the double helix each acting as a template for the replication of the other strand. Not all chemical changes to DNA molecules are preserved through DNA replication however (changes to the backbone are not maintained but changes in the sequence of bases are). Moreover, mechanisms of proofreading and repair have evolved to correct ‘errors’ of replication (Sterelny et al. 1996). But some changes remain uncorrected, as are subsequent changes to these changes, allowing the exploration of a vast space of possible sequences. Muller recognized that this open-ended property of genetic change had far-reaching consequences:

Thus it is not inheritance and variation which bring about evolution, but the inheritance of variation, and this in turn is due to the general principle of gene construction which causes the persistence of autocatalysis despite the alteration in the structure of the gene itself. Given, now, any material or collection of materials having this one unusual characteristic, and evolution would automatically follow, for this material would, after a time, through the

⁸ Here I employ a loose definition of replication and heritability as recurrence of form, including the ‘simple replicators’, ‘limited hereditary replicators’, and ‘unlimited hereditary replicators’ of Maynard Smith and Szathmáry (1995, pp. 41–42).

accumulation, competition and selective spreading of the self-propagated variations, come to differ from ordinary inorganic matter in innumerable respects, in addition to the original difference in its mode of catalysis. There would thus result a wide gap between this matter and other matter, which would keep growing wider, with the increased complexity, diversity and so-called 'adaptation' of the selected mutable material. (page 35)

Not all things that are replicated have the property that changes in their structure, arising by chance and selected by the environment, are transmitted to future generations without compromising autocatalysis. And even if some non-genetic replicators transmit minor changes in this manner, few, if any, have potential for the open-ended adaptive change that is characteristic of DNA sequences. Human cultural evolution clearly has this open-ended quality (Boyd et al. 2011) although the nature of the heritable difference-makers of cultural change, if such exist, is disputed. One needs to have evolved very sophisticated organisms by other processes before there can be meaningful cultural evolution. Genes are special (and so, in its own way, is culture).

The strategic gene is a refinement of the metaphor of genes as self-interested agents. The phenotypes that are chosen by natural selection resemble those that would be chosen by a rational agent attempting to ensure its own transmission to future generations. Mindless genes can therefore be viewed *as if* they make strategic decisions. Some find this metaphor appealing (Queller 2011; Dennett 2011) whereas others consider it insidious and symptomatic of paranoia (Godfrey-Smith 2009, p. 144). Agential metaphors are less attractive (perhaps less seductive) for most other hereditary replicators, such as DNA methylation, membranes, nests, or money invested in the stock market. Are gene-selectionists inconsistent or is there some principled difference between genes and non-agential replicators? I believe the difference resides in the sophistication of genes' strategies made possible by the open-ended nature of genetic inheritance identified by Muller. Genes are 'indefinite hereditary replicators' (Maynard Smith and Szathmáry 1995, p. 58) that accumulate functional information about what works in the environment to a much greater extent than other candidates for the replicator role.

Acknowledgments Daniel Dennett, Steve Frank, Andy Gardner, Peter Godfrey-Smith, Kim Sterelny and the Fundamental Interconnectedness of All Things discussion group made helpful comments on the manuscript.

References

- Bacolla A, Ulrich MJ, Larson JE, Ley TJ, Wells RD (1995) An intramolecular triplex in the human γ -globin 5'-flanking region is altered by point mutations associated with hereditary persistence of fetal hemoglobin. *J Biol Chem* 270:24556–24563
- Bergstrom CT, Rosvall M (2011) The transmission sense of information. *Biol Philos* 26:159–176
- Bouchard F, Rosenberg A (2004) Fitness, probability and the principles of natural selection. *Br J Philos Sci* 55:693–712
- Boyd R, Richerson PJ, Henrich J (2011) The cultural niche: why social learning is essential for human adaptation. *Proc Natl Acad Sci USA* 108:10918–10925
- Dawkins R (1976) *The selfish gene*. Oxford University Press, Oxford

- Dawkins R (1982) *The extended phenotype*. Oxford University Press, Oxford
- Dennett DC (2011) *Homunculi rule: reflections on Darwinian populations and natural selection* by Peter Godfrey-Smith. *Biol Philos* 26:475–488
- Ewens WJ (2011) What is the gene trying to do? *Br J Philos Sci* 62:155–176
- Fisher RA (1941) Average excess and average effect of a gene substitution. *Ann Eugen* 11:53–63
- Fisher RA (1958) *The genetical theory of natural selection*, 2nd edn. Dover, New York
- Frank SA (2009) Natural selection maximizes fisher information. *J Evol Biol* 22:231–244
- Gardner A, Welch JJ (2011) A formal theory of the selfish gene. *J Evol Biol* 24:1801–1813
- Godfrey-Smith P (2009) *Darwinian populations and natural selection*. Oxford University Press, Oxford
- Gray R (1992) Death of the gene: developmental systems strike back. In: Griffiths P (ed) *Trees of life: essays in philosophy of biology*. Kluwer, Dordrecht, pp 165–209
- Griffiths PE (1998) What is the developmentalist challenge? *Philos Sci* 65:253–258
- Haig D (1997) The social gene. In: Krebs JR, Davies NB (eds) *Behavioural ecology*, 4th edn. Blackwell Scientific, Oxford, pp 284–304
- Haig D (2003) Meditations on birthweight: Is it better to reduce the variance or increase the mean? *Epidemiology* 14:490–492 (erratum *Epidemiology* 14:632)
- Haig D (2006) The gene meme. In: Grafen A, Ridley M (eds) *Richard dawkins. How a scientist changed the way we think*. Oxford University Press, Oxford, pp 50–65
- Haig D (2007) Weismann rules! OK? Epigenetics and the Lamarckian temptation. *Biol Philos* 22:415–428
- Haig D (2008) Huddling: brown fat, genomic imprinting, and the warm inner glow. *Curr Biol* 18:R172–R174
- Haig D (2010) The huddler's dilemma: a cold shoulder or a warm inner glow. In: Székely T, Moore AJ, Komdeur J (eds) *Social behaviour: genes, ecology and evolution*. Cambridge University Press, Cambridge, pp 107–109
- Haig D, Trivers R (1995) The evolution of parental imprinting: a review of hypotheses. In: Ohlsson R, Hall K, Ritzen M (eds) *Genomic imprinting: causes and consequences*. Cambridge University Press, Cambridge, pp 17–28
- Lewontin RC (1974) The analysis of variance and the analysis of causes. *Am J Hum Genet* 26:400–411
- Lewontin RC (2000) Foreword. In: Oyama S (ed) *The ontogeny of information*, 2nd edn. Duke University Press, Durham, pp vii–xv
- Maynard Smith J, Szathmáry E (1995) *The major transitions in evolution*. W. H. Freeman, Oxford
- Muller HJ (1934) Variation due to change in the individual gene. *Am Nat* 56:32–50
- Okasha S (2008) Fisher's fundamental theorem of natural selection—a philosophical analysis. *Br J Philos Sci* 59:319–351
- Olson M (1961) *The logic of collective action. Public goods and the theory of groups*. Harvard University Press, Cambridge
- Oyama S (2000) *The ontogeny of information. Developmental systems and evolution*, 2nd edn. Duke University Press, Durham
- Queller DC (2011) A gene's eye view of Darwinian populations. *Biol Philos* 26:905–913
- Sandhu KS, Shi C, Sjölander M, Zhao Z, Göndör A, Liu L et al (2009) Nonallelic transvection of multiple imprinted loci is organized by the *H19* imprinting control region during germline development. *Genes Dev* 23:2598–2603
- Shea N (2007) Representation in the genome and in other inheritance systems. *Biol Philos* 22:313–331
- Sober E (1984) *The nature of selection*. MIT Press, Cambridge
- Sober E, Wilson DS (1994) A critical review of philosophical work on the units of selection problem. *Philos Sci* 61:534–555
- Sterelny K, Griffiths PE (1999) *Sex and death*. Chicago University Press, Chicago
- Sterelny K, Kitcher P (1988) The return of the gene. *J Philos* 85:339–361
- Sterelny K, Smith KC, Dickison M (1996) The extended replicator. *Biol Philos* 11:377–403
- Sturtevant AH (1915) The behavior of the chromosomes as studied through linkage. *Z induct Abstam Vererbungs* 13:234–287
- Williams GC (1966) *Adaptation and natural selection*. Princeton University Press, Princeton
- Williams GC (1986) Comments on Sober's *The nature of selection*. *Biol Philos* 1:114–122
- Wilson DS, Sober E (1994) Reintroducing group selection to the human behavioral sciences. *Behav Brain Sci* 17:585–654
- Wunner WH (2007) Rabies virus. In: Jackson AC, Wunner WH (eds) *Rabies*, 2nd edn. Elsevier, Amsterdam, pp 23–68