The "Genetic Program" Program: A Commentary on Maynard Smith on Information in Biology

Kim Sterelny

Philosophy of Science, Vol. 67, No. 2. (Jun., 2000), pp. 195-201.

Stable URL:
http://links.jstor.org/sici?sici=0031-8248%28200006%2967%3A2%3C195%3AT%22PPAC%3E2.0.CO%3B2-L

Philosophy of Science is currently published by The University of Chicago Press.

The "Genetic Program" Program: A Commentary on Maynard Smith on Information in Biology*

Kim Sterelny†‡
Department of Philosophy, Victoria University of Wellington
Philosophy Program, RSSS, Australian National University

1. Common Ground and an Agenda. In many texts on evolution the reader will find a characteristic depiction of inheritance and evolution, one showing the generations of an evolving population linked only by a causal flow from genotype to genotype. On this view, the genotype of each organism in this population plays a dual role as both the motor of individual development and as the sole causal channel across the generations (cf., e.g., Maynard Smith 1993, Fig. 8). This picture is known to be literally false. In many species, parents exert direct causal influence on their offspring, and some of those influences cause the offspring to resemble the parent. For example, a butterfly that lays her eggs on the same plant host on which she hatched thereby exerts a causal influence on her offspring, and one apt to cause them to resemble her more than they would, had she chosen a plant of a different species. None of this is at all controversial, but it poses a puzzle. If the "Weissmanian" conception is literally false, why is it seen as a perspicacious representation of evolution?

Two main answers are on offer. Richard Dawkins has argued that the genetic channel between the generations is the only one of evolutionary significance. It is the only channel that contributes to cumulative evolutionary change, for genes and (almost) only genes meet the *replicator condition*. If there is a change in a gene copied from the parent to the F1 generation, that change will be copied through to the F2 and subsequent generations (Dawkins 1982a, 97-98). With very limited exceptions, that is

*Received January 2000; revised February 2000.

†Send requests for reprints to the author, Department of Philosophy, Victoria University of Wellington, P.O. Box 600, Wellington, New Zealand.

‡Thanks to Russell Gray and Paul Griffiths for comments on a draft of these comments.

Philosophy of Science, 67 (June 2000) pp. 195–201. 0031-8248/2000/6702-0002$2.00

not true of other causal influences across the generations. The second option is to appeal to the special developmental role of the genes. Only genes carry information. If development just is the exploitation of genetic information, then evolution just is change in this developmental information over time. Maynard Smith defends the second of these strategies. In doing so, he describes himself as giving a "natural history" rather than an analysis of information. So in this commentary I develop a Maynard Smithian analysis to reveal some surprising commitments of his program.

In the debate on genetic information, it is common ground that the standard apparatus of information theory applies to the phenotype/genotype relation, but that information so understood is too weak a notion to sustain his program. The standard apparatus defines information as the covariation between a signal and its source. Holding environmental factors constant, genotypes covary with phenotypes. But other factors causally relevant to development also carry predictive information. The plant on which the butterfly eggs are laid covaries with developmental outcomes. So genes predict phenotypes, but they are not alone in doing so. Moreover, if information is just covariation, information cannot be misread or misused. In certain developmental contexts, typical human genotypes predict that the resulting phenotype has vestigial limbs. That is, in those contexts they covary with that outcome. So if we are speaking of *covariational information*, thalidomide does not cause the genetic program to be misread. It is not in dispute that genes carry covariational information, but Maynard Smith's program needs a richer, semantic, notion, one that is strong enough to support the idea of error and misrepresentation. And unlike the covariational notion, it may mark genes as playing a special role in development.

In order to defend the conception of the genome as a genetic program, Maynard Smith must solve the following two problems. First, he needs a semantic conception of information that applies to the gene/phenotype relationship and, if Weissmanism is to be defended, only to that relationship (perhaps with limited exceptions). Second, if the information coded in the genotype explains development, it must be read and used. It is a well-known feature of the covariational conception of information that the information/channel condition distinction is a fact about us and our interests rather than a fact about the world. For a geneticist, the *eyeless* gene's relation to eye formation is information, and the receptors in the cell that respond to the protein that it makes are channel conditions. For a cell biologist the situation may be reversed: the gene/protein system is the channel condition that reveals the structure of the protein-using cell mechanism. But if development is driven by information decoded from the genome, the information/reader distinction must be objective and robust rather than interest relative.

2. Teleoinformation. Where does genetic information come from? Maynard Smith develops the only live option. The genotype carries semantic information in virtue of its biological function. A particular gene consists in a particular sequence of bases in a particular reading frame because such sequences have lead to certain phenotypic outcomes, and the sequence has been maintained in the population through selection for that outcome. The gene is not merely correlated with a trait: that trait explains why that gene has its form. Hence when a human genotype results in a phenotype with dwarfed arms, the information in it has been misread. The actual phenotype varies critically from the phenotype that mediated the selection of the gene.

This is an idea that has been floated before (Sterelny, Smith, and Dickison 1996; Sterelny and Griffiths 1999). One apparent consequence is that all adapted developmental resources turn out to carry semantic information about phenotypic outcomes. Thus if the cytoplasmic factors in the cell that turn on genes in the early embryo have their chemical makeup because they have been selected to serve that function, then those cytoplasmic factors carry semantic information about the basic morphological layout of the embryo. Obligatory symbiosis provides other examples of adapted developmental resources. Many arthropods require microorganisms to produce nutrients on which they depend for growth. Such microorganisms are transmitted in the egg, and the mechanisms mediating symbiont transfer can be very precise (see Morgan and Baumann 1994, Frank 1996). So, symbionts are not just physically transmitted across host generations. They are adapted developmental resources: their coevolved adaptations encode semantic information about host phenotypes. Genetic program theorists often assume that symbiont transfer is an epiphenomenon of gene flow, and hence do not take the apparent extension of the semantic conception of information to other adapted developmental resources seriously. So, it is important to see that this assumption is question-begging. It presupposes that the development of the organism, including its extended phenotype, is just the execution of its genetic program. And it makes an untested empirical assumption: all heritable differences in symbiont transfer within a population covary with, and are tracked by, genetic differences (Gray 1992).

Maynard Smith's model has resources that allow him to avoid concluding that all adapted developmental resources carry semantic information about phenotypes. For he marries the appeal to biological function to a second criterion. The relationship between gene and trait is arbitrary. His example is of an inducer gene, which switches off a repressor protein made by a regulatory gene. Any protein that would bind to the repressor protein to alter its shape would do, and many proteins would do that. The inducer gene is constrained by its function, but it is not determined by it.

To that extent, the relationship between gene and its functional outcome—the phenotype it codes for—is arbitrary.[1]

It is arguable that the relationship between a symbiotic microorganism and its phenotypic effect on a host is not arbitrary. For microorganisms are recruited and re-engineered for symbiotic roles precisely because they were pre-adapted in their independent lives for these effects on hosts. It is no accident that it is a fungus that has been recruited as the leafcutter's partner. Fungi do not need light, so they can be grown within the nest. Many can reproduce vegetatively, so the complexities of spore germination can be bypassed. More generally, as I have argued elsewhere (Sterelny 2000), there is an important distinction between sample-based inheritance systems and others. Transmitting the critical resource itself, or a sample of it, across the generation allows one to dispense with information. Since the developing leafcutter colony has a sample of the fungus to grow, it does not need information on how to make the fungus.

So perhaps some developmental resources do not carry semantic information about the developed phenotype despite having developmental biofunctions. For they are part of sample-based systems of inheritance and development, rather than code-based, arbitrary systems. Since the evolutionary dynamics and evolutionary significance of sample-based systems may well be different from arbitrary systems (Sterelny 2000), this distinction is not ad hoc. But what of cell cytoplasm? The cytoplasm in a cell is clearly inherited from the mother: it is part of cross-generation causal flow. And it is arbitrary. It is presumably a contingent fact of biology, rather than a necessary fact of biochemistry, that certain chemical gradients switch on a small set of embryonic genes, thus determining the basic back to front, up and down, and right and left organization of the body. Maynard Smith argues with respect to the *lac* gene that gene on/off switches are arbitrary. So he is committed to the view that cytoplasmic factors in the early developing egg have their adapted effect on the phenotype contingently. Is he thereby committed to the view that the cytoplasm of the fertilized egg is part of the developmental program? Not quite. For perhaps genes, but not cytoplasm gradients, depend on mechanisms which read and use information. The next section unpacks this idea.

3. Executing the Genetic Program. We seem to have developed the following picture. Within some lineages, amongst the total set of causal factors

1. Not perhaps as arbitrary as language, where referent does not even constrain a term, but as arbitrary as icons and the like. Such symbols are constrained by their function—not any shape at all would do to indicate the male toilet—but function does not fix shape. Seeing it this way suggests that the criterion defines a continuum rather than a dichotomy.

regularly contributing to the development of an organism, some exist, and exist in that form, because they are adapted for a specific developmental role. Among the adapted developmental resources that appear generation by generation, some are contingently connected to their phenotypic product, and others are more or less intrinsically connected to that product. That is, given that they are recruited as developmental resources at all, their adaptive effect on the phenotype is predictable. Finally, of those that are contingently connected to their phenotypic product, some depend on readers for their phenotypic upshot. For if the genotype is a developmental program then there must be mechanisms that read and execute that program. Information is causally inert unless it is recognized and used. So perhaps genes, but not cytoplasmic chemical gradients, are read.

In particular, I think Maynard Smith should see regulatory genes as gene readers rather than as encoding information in their own right. The mouse *eyeless* gene does not carry information about how to make a mouse eye. It reads this information off other genes. For when the mouse *eyeless* gene is transplanted to a fruitfly and activated, the result is an ectopic *fruitfly* eye. A fully developed compound eye will form on a fruitfly leg. This suggests that *eyeless* is a reader capable of reading many messages. In mice, its executes (with, of course, much help from the surrounding milieu) mouse eye-making instructions; in fruitflies, it executes compound eye-making instructions. Perhaps in other organisms it would make a fish eye, or a simple pinhole eye. Maynard Smith talks of *eyeless* as signalling "make an eye here," and I suspect he would resist the view that regulatory genes are readers. But that is the most consistent development of his argument, and it connects to a major theme of his other work.

The idea in play is that regulatory genes within a given domain systematically map *differences* in structural genes onto *differences* in phenotypes. Hence they are channel conditions. If this is right, then we can connect the "genetic program" program to the evolution of unlimited rather than limited systems of heredity. Maynard Smith has argued that only unlimited systems make cumulative selection possible, hence only unlimited systems make complex adaptation possible (Szathmáry and Maynard Smith 1997). There is a vast number of DNA sequences, but that by itself does not constitute an unlimited system of heredity. For that, sequence differences have to be reliably mapped onto phenotypic differences. If most of these differences were phenotypically silent, or if their phenotypic upshot varied unpredictably from generation to generation, then DNA transmission would not constitute an unlimited system of heredity. But if there are readers that are systematically sensitive to the differences between sequences, then DNA transmission would indeed be an unlimited system. There is a short-range reading system—the gene/protein code systematically maps sequence differences onto protein differences. But it does not

follow that genes code for phenotypes rather than just proteins (Godfrey-Smith 1999, forthcoming). However, the example of *eyeless* suggests that regulatory genes are part of a gene reading system that has systematic and predictable effects at the level of whole phenotypes, though of course in context dependent ways.

4. Upshot. In sum, on this view some but not all genes carry developmental information because: (i) they are adapted for their roles in development and evolution; (ii) they are part of an inheritance system each element of which has a contingent connection to the developmental outcome it programs; (iii) they are part of an unlimited system of heredity, a system that is unlimited because other elements in the developmental system have the job of information readers. Within specific contexts and with limited reliability, they turn sequence differences into phenotypic differences.

If this is the right picture to extract from Maynard Smith's natural history of information, it does not just vindicate the current heuristic use of informational notions—it defines a critical empirical project. Can we actually find discrete, fairly autonomous physical mechanisms that are capable of mapping genotype differences onto phenotype differences? The mechanisms that map genotype differences onto protein differences have been identified, but we cannot just assume that such mechanisms must exist for the phenotype as a whole. A parallel with connectionist models in cognitive science is instructive at this point. They show we cannot assume that a system capable of generating a complex outcome must divide into elements which carry information and elements which recognize and use that information. Connectionist architectures blur the distinction between data structure and programs that use that data. The outcome of development is both complex and predictable. But that does not in itself guarantee that there must exist discrete physical systems that read information and other discrete structures that encode it. Developmental systems theorists are sceptical about informational views of the gene because they see no evidence within development for this difference (Gray 1992, Griffiths and Gray 1994).

Furthermore, grant that genetic information is real. The mapping between sequence and phenotype is stable, rich, and mediated by channels adapted to use that information. Is it the only such system? Consider once again chemical gradients in cell cytoplasm. Would small differences in gradients be "read"? That is, would small differences in cell cytoplasm result in small differences in morphological organization; not always, and in any circumstances, but often enough in the actual developmental contexts common in the relevant lineage? We do not know: the range of variation generated by non-genetic inheritance remains an open and understudied question (Gray forthcoming).

In summary, I doubt that the only mechanism of inheritance is gene flow. But though there are several inheritance systems, their characteristics are importantly different. Some inheritance systems but not others may transmit semantic information between the generations. That remains a live hypothesis. But unpacking the commitments of that hypothesis is no trivial task, and to the extent that the hypothesis is clear, it seems to me that neither all nor only genes are information carriers.

REFERENCES

Dawkins, R. (1982a), *The Extended Phenotype*. Oxford, Oxford University Press.

Frank, S. A. (1996), "Host Control of Symbiont Transmission: The Separation of Symbionts into Germ and Soma", *American Naturalist* 148: 1113–1124.

Godfrey-Smith, P. (1999), "Genes and Codes: Lessons from the Philosophy of Mind?", in V. Hardcastle (ed.), *Biology Meets Psychology: Constraints, Conjectures, Connections*. Cambridge, MA: MIT Press, 305–322.

———. (forthcoming), "On The Theoretical Role of Genetic Coding", *Philosophy of Science*.

Gray, R. (1992), "Death of the Gene: Developmental Systems Strike Back", in P. Griffiths (ed.), *Trees of Life: Essays of the Philosophy of Biology*. Dordrecht, Kluwer: 165–209.

———. (forthcoming), "Selfish Genes or Developmental Systems? Evolution Without Replicators and Vehicles", in R. Singh, K. Krimbas, D. Paul, and J. Beatty (eds.), *Thinking About Evolution: Historical, Philosophical and Political Perspectives: Festschrift for Richard Lewontin*. Cambridge: Cambridge University Press.

Griffiths, P. E. and R. Gray (1994), "Developmental Systems and Evolutionary Explanation", *Journal of Philosophy* XCI(6): 277–304.

Maynard Smith, J. (1993), *The Theory of Evolution* [1st ed., 1975]. Cambridge: Cambridge University Press.

Sterelny, K. (2000), "Niche Construction, Developmental Systems and the Extended Replicator", in R. Gray, P. Griffiths and S. Oyama (eds.), *Cycles of Contingency*. Cambridge, MA: MIT Press.

Sterelny, K. and P. Griffiths (1999), *Sex and Death: An Introduction to Philosophy of Biology*. Chicago: University of Chicago Press.

Sterelny, K., K. Smith, and M. Dickison (1996), "The Extended Replicator", *Biology and Philosophy* 11: 377–403.

Szathmáry, E. and J. Maynard Smith (1997), "From Replicators to Reproducers: The First Major Transitions Leading to Life", *Journal of Theoretical Biology* 187: 555–571.

References

Host Control of Symbiont Transmission: The Separation of Symbionts Into Germ and Soma

Steven A. Frank

The American Naturalist, Vol. 148, No. 6. (Dec., 1996), pp. 1113-1124.

Stable URL:

http://links.jstor.org/sici?sici=0003-0147%28199612%29148%3A6%3C1113%3AHCOSTT%3E2.0.CO%3B2-5

Developmental Systems and Evolutionary Explanation

P. E. Griffiths; R. D. Gray

The Journal of Philosophy, Vol. 91, No. 6. (Jun., 1994), pp. 277-304.

Stable URL:

http://links.jstor.org/sici?sici=0022-362X%28199406%2991%3A6%3C277%3ADSAEE%3E2.0.CO%3B2-%23