Genetic Information: A Metaphor In Search of a Theory*

Paper submitted to *Philosophy of Science*

Paul E. Griffiths
Department of History and Philosophy of Science,
1017 Cathedral of Learning,
University of Pittsburgh,
Pittsburgh, PA 15260 USA

Email: pauleg@pitt.edu

Abstract

John Maynard Smith has defended against philosophical criticism the view that developmental biology is the study of the expression of information encoded in the genes by natural selection. However, like other naturalistic concepts of information, this 'teleosemantic' information applies to many non-genetic factors in development. Maynard Smith also fails to show that developmental biology is concerned with teleosemantic information. Some other ways to support Maynard Smith's conclusion are considered. It is argued that on any definition of information the view that development is the expression of genetic information is misleading. Some reasons for the popularity of that view are suggested.

1. Introduction: Information Talk in Biology

> An organism's physiology and behaviour are dictated largely by its genes. And those genes are merely repositories of information written in a surprisingly similar manner to the one that computer scientists have devised for the storage and transmission of other information... (Economist 1999, 97)

The only controversial feature of this statement is the claim that behavior is largely genetic. It is conventional wisdom that insofar as the traits of an organism are subject to biological explanation, those traits express information coded in the organism's genes. Conventional wisdom has recently been defended in this journal by the eminent biologist John Maynard Smith (Maynard Smith 2000). I will argue, however, that the only truth reflected in the conventional view is that there is a genetic code by which the sequence of DNA bases in the coding regions of a gene corresponds to the sequence of amino acids in the primary structure of one or more proteins. The rest of 'information talk' in biology is no more than a picturesque way to talk about correlation and causation. The claim that biology 'is, itself, an information technology' (Economist 1999, 97) is on a par with the claim that the planets compute their orbits around the sun. Taking 'information talk' in biology too seriously is not merely a journalist's error. Many biologists, when asked to talk about their discipline in broad, philosophical terms, would represent it in the same light. Nevertheless, as Sahotra Sarkar has noted:

> there is no clear, technical notion of "information" in molecular biology. It is little more than a metaphor that masquerades as a theoretical concept and ...leads to a misleading picture of possible explanations in molecular biology. (Sarkar 1996,187)

Correcting the conventional wisdom is important in its own right, but it also clears the ground for an important, new approach to the debate over 'genetic determinism'. Phillip Kitcher has joked that arguing against genetic determinism is like battling the undead (Kitcher In Press). The 'interactionist consensus' that all traits depend on both genetic and environmental factors has been reaffirmed again and again, but the discovery that genes are involved in the development of a trait continues to be used to infer that the trait will be hard to change by manipulating the environment. This inference persists in the face of widespread agreement that there are always other causal factors involved in the development of 'genetic' traits and many examples of 'genetic' traits being modified via these factors. For example, genetic diseases can be treated by environmental (i.e: drug or dietary) interventions. "But why, then," Kitcher's frustrated interactionist asks, "do we always end up discussing whether genotypes are all-powerful in development?" (Kitcher In Press). In my view, an important part of the answer has been given by Susan Oyama (Oyama 2000a, 2000b). Genetic causation is interpreted deterministically because genes are thought to be a special kind of cause. Genes are instructions - they provide information - whilst other causal factors are merely material. The intuitive notion of information is a semantic notion, carrying the implication that genes, unlike other causal factors, are about, or directed at, the outcomes they help to produce. Little wonder, then, that the gene-trait relationship seems intuitively more context-independent than the relationship between traits and other causes. The temperature of the nest determines the sex of an crocodile, the eggs developing as male at intermediate temperature and female at extreme temperatures, but outside that very particular context the cause – a mound of

rotting vegetation with a temperature between *n* and *m* degrees - retains no connection with masculinity. In contrast, even when the informational gene fails to produce its effect, it remains directed at that effect. A 'gay gene' is an instruction to be gay even when the person carrying it is straight.

If Oyama is correct, then finding a new way to think about genetic causation is a necessary prerequisite to laying the unquiet ghost of the nature/nurture controversy and learning to think clearly about the interaction of genetic and other factors in development. It is this prospect that makes it important to bring home the fact that 'information talk' in biology is merely picturesque. But the deflation of information talk in biology is logically independent of Oyama's larger claim about the causes of genetic determinism. Kitcher, for example, accepts a deflationary account of information talk, but rejects Oyama's claim that loose information talk sustains genetic determinism (Kitcher In Press). The position I will defend here is not the entire Oyama package - the 'developmental systems theory' - but only what I have elsewhere called the 'parity thesis' (Griffiths and Knight 1998). Any defensible definition of information in developmental biology is equally applicable to genetic and non-genetic causal factors in development. Definitions of information on which genes contain developmental information but methylation patterns or incubation temperatures do not are illegitimate because they cannot be 'naturalised' - they ascribe properties to genes that cannot be grounded in physical and biological facts. Hence, if discarding information talk would be useful, there is no substantial, biological reason not to discard it.

2. Two Concepts of Information

Concepts of information can be divided into two rough categories - *causal* information concepts and *intentional* information concepts. Causal conceptions of information derive from the mathematical theory of communication (Shannon and Weaver 1949). Mathematical information theory studies only the quantity of information in a physical system. The quantity of information in a system can be understood roughly as the amount of order in that system, or the inverse of the entropy (disorder) that all closed physical systems accumulate over time. This measure says nothing about the content of information. However, there is a related causal notion of information content (Dretske 1981). Information flows over a channel connecting two systems, a receiver that contains the information and a sender, the system that the information is about. There is a channel between two systems when the state of one is systematically causally related to the other, so that the state of the sender can be discovered by observing the state of the reciever. The causal information content of a signal is simply the state of affairs with which it reliably correlated at the other end of the channel. Thus, smoke carries information about fire and disease phenotypes carry information about disease genes.

The second category of information concept deals with intentional information or 'semantic information' (Godfrey-Smith 1999). It is information in this sense that human thoughts and utterances are supposed to contain. If there is a relationship between intentional information and causal information it is a complex and distant one. We think about things with which we have only the most tenuous causal connection (e.g: undiscovered galaxies) and about things that do not exist (e.g: phlogiston or Pope Joan).

The relation between thoughts and these various objects of thought is 'intentionality' and the question of how physical systems like brains can exhibit intentionality is one of the most vexed issues in the philosophy of mind. One of the distinctive features of intentional information is that it can be false - it can misrepresent how things are (Godfrey-Smith 1989). It is difficult to reproduce the phenomena of misrepresentation using a causal notion of information. A signal cannot both correlate with a source and not correlate with it, nor can a signal correlate with a source that does not exist. The most promising attempts to give a naturalistic account of intentional information are the so-called 'teleosemantic' theories to be discussed below, according to which a sign represents whatever evolution designed it to represent. Genetic information is usually described as if misinterpretation made sense and hence appears to be intentional information. For instance, it is a commonplace to say that a genome contains a 'disjunctive program' or 'open program'. The genetic message takes the form 'develop like *this* under these circumstances, like *that* under other circumstances'. Evolutionary psychologists seek to explain cultural differences between human minds as just this kind of disjunctive response of the human developmental program to various environmental factors (Cosmides, Tooby et al. 1992). In contrast, no one says that the human genome encodes the instruction 'when exposed to the drug thalidomide grow only rudimentary limbs'. This would be one branch of the disjunctive program if we were talking about the causal information in the human genome. When the relevant channel is contaminated by thalidomide, human genes send this causal information. The fact that the notion of a disjunctive program is not applied to outcomes that are thought to be pathological or

accidental suggests that the information in the program is usually conceived of as intentional information.

3. Maynard Smith on Genetic Information

The obvious way to defend information talk in biology is to argue that it is just like the many other, scientifically respectable uses of causal information. This has been the traditional strategy. In an address to the Fifth International Congress of Evolutionary and Systematic Biology in 1996, John Maynard-Smith made use of this conventional defense. Information talk in biology is to be interpreted 'more or less in the spirit of information theory' (author's notes). The disadvantage of this defense is that it implies parity between genetic and non-genetic causes in development. Information in the sense legitimated by information theory is the systematic dependence of a signal on a source, a dependence that is created by a set of channel conditions. In the case of development, the genes are normally taken to be the source, the life-cycle of the organism is the signal and the channel conditions are all the other resources needed for the life-cycle to unfold. But it is a fundamental feature of information theory that the role of source and channel condition can be reversed. The old television 'test-card' did exactly that, holding the transmission constant so that the television engineer could read off the state of what were previously channel conditions. The source/channel distinction is imposed on a natural causal system by the observer. A source is simply one channel condition whose current state the signal is being used to investigate. If all other resources are held constant, a life-cycle can give us information about the genes, but if the genes are held constant, a life-cycle can give us information about whichever other resource we decided to let vary. So far as causal

information goes, every resource whose state affects development is a source of developmental information (Johnston 1987; Gray 1992; Griffiths and Gray 1994; Oyama 2000a).

The fact that causal information conforms to this 'parity thesis' is now quite widely recognized (Godfrey-Smith 1999; Sterelny and Griffiths 1999; Maynard Smith 2000; Kitcher In Press). In his most recent discussion of causal information, Maynard Smith notes that:

> With this definition, there is no difficulty in saying that a gene carries information about adult form; an individual with the gene for achondroplasia will have short arms and legs. But we can equally well say that a baby's environment carries information about growth; if it is malnourished, it will be underweight. (Maynard Smith 2000, 189)

In response to the threat of parity, Maynard Smith has abandoned the idea that biological information is causal information. He continues:

> informational language has been used to characterize genetic as opposed to environmental causes. I want now to try to justify this usage.
>
> I will argue that the distinction can be justified only if the concept of information is used in biology only for causes that have the property of intentionality.... A DNA molecule has a particular sequence because it specifies a particular protein, but a cloud is not black because it predicts rain. This element of intentionality comes from natural selection. (Maynard Smith 2000, 189-90)

Maynard Smith proposes to analyse the intentionality of genetic information using *teleosemantics*, the philosophical program of reducing meaning to biological function (teleology) and then reducing teleology to natural selection (Millikan 1984; Papineau 1987). Whilst there is considerable controversy about whether such reductions can be successfully carried out, teleosemantics remains one of the more popular programs for naturalizing intentionality. Maynard Smith begins by arguing that natural selection is strongly analogous to computer programming using the 'genetic algorithm' technique. The genetic algorithm programmer randomly varies the code of a computer program and selects variants for their performance. In the same way, natural selection randomly varies the genes of organisms and selects those organisms for their fitness. Just as the function of the selected computer program is to perform the task for which it was selected, the biological function of successful genes is to produce the developmental outcomes in virtue of which they were selected. Such genes are intentionally directed onto, or about, those effects. The defective haemoglobin gene in some human populations, which has been selected because it sometimes confers resistance to malaria, carries teleosemantic information about malaria resistance.

However, teleosemantic information is fundamentally unsuited to Maynard-Smith's aim of avoiding parity. The most developed version of the teleosemantic theory of genetic information is Sterelny et al's 'extended replicator theory' (Sterelny, Dickison et al. 1996). Sterelny and his collaborators recognize from the outset that teleosemantic information exists in both genetic and in some non-genetic developmental causes. Sterelny reiterates this view in his reply to Maynard Smith (Sterelny 2000). Russell Gray and I have argued

that teleosemantic information exists in a very much wider range of developmental causes that Sterelny et al suggest (Griffiths and Gray 1997), but my argument here follows even on the more conservative view. Teleosemantic information exists in any inheritance system that is a product of evolution, including epigenetic inheritance systems. The term 'epigenetic inheritance system' is used to denote any biological mechanisms which produces resemblances between parents and offspring and which works in parallel with the inheritance of nuclear and mitochondrial (Jablonka and Lamb 1995; Jablonka and Szathmáry 1995). Every organism inherits a great deal besides its DNA. To develop normally the egg cell must contain features such as: basal bodies and microtubule organising centres, correct cytoplasmic chemical gradients, DNA methylation patterns, membranes and organelles, as well as DNA. Changes in these other resources can cause heritable variation which appears in all the cells descended from that egg cell. One of the best-understood of these mechanisms is the DNA methylation inheritance system. A methylation pattern is a series of additional chemical groups attached to a DNA sequence. Methylation patterns block transcription of any genes they cover and they are replicated by the methylation copying system in all the cells descended from a given cell. Differences in methylation are important in tissue differentiation during the lifetime of a single organism, but they can also pass between the generations. Methylation patterns are often applied to the DNA in a sperm or egg by the parent organism. DNA methylation inheritance has excited a great deal of interest because of it is easy to see how it could play a role in conventional, micro-evolutionary change. A typical example comes from some controversial British research on behavioral differences in male and female children. The proposed mechanism for the transmission of

these behavioral differences is that female humans methylate a sequence of the X-chromosome in their gametes, so that males, who get only one X-chromosome and get it from their mother, cannot transcribe the genes in that region. Hence certain gene products are denied to all males. Males demethylate that sequence in their sperm cells, so females get readable copies of these genes on the X-chromosomes they receive from their fathers (Skuse and al 1997). Naturally, when this research was reported in the media it was announced as the discovery that sex differences are in the genes: 'Genes say boys will be boys, and girls will be sensitive' (Radford 1997).

Epigenetic inheritance research has focused on mechanisms within the cell. Developmental systems thinkers have emphasized a wider range of epigenetic mechanisms (Gray 1992; Griffiths and Gray 1994; Griffiths and Gray 2001). The characteristics of epigenetic inheritance systems within the cell are shared by many extra-cellular structures. Some castes of the aphid *Colophina arma* require a growth spurt as part of their life-cycle. These and only these castes inherit the micro-organisms that make the chemicals on which this growth spurt depends (Moran and Baumann 1994). The morphology of queens and the colony structure of the fire ant *Solenposis invicta* differ radically between genetically similar lineages of the species because of stably replicated pheremonal nest 'cultures' (Keller and Ross 1993). Any queen raised in a colony with a particular culture will found a colony with the same culture, as shown by moving eggs from one culture to the other. Many parasites, both vertebrate and invertebrate, maintain associations with particular host species over evolutionary time through "host-imprinting". Thus, some insects lay their eggs on the plant whose leaves they tasted as

larvae. Some parasitic finches lay their eggs in the nest of the host species which they imprinted on as chicks (Immelmann 1975). “Host switching" occurs on the rare occasions when this mechanism malfunctions and the mother lays her eggs on the wrong plant or in the wrong nest. On the still more rare occasions when these misplaced eggs flourish, the imprinting mechanism will ensure that this new, epigenetic mutant form reliably reproduces itself. Clearly, all the mechanisms discussed here are candidates for evolutionary explanation - they did not come about by accident. This means that the physical traces by which these mechanisms operate have biological functions and thus, on the teleosemantic approach recommended by Maynard Smith, that these traces contain information. Hence, just like causal information, teleosemantic information obeys the parity thesis.

The mechanism of epigenetic inheritance that takes us furthest from the nucleus is so-called 'niche construction' (Odling-Smee 1988; Odling-Smee, Laland et al. 1996; Laland, Odling-Smee et al. 2001). Many features of an organism's niche exist only because of the effects of previous generations of that species on the local environment. One of the earliest examples of this phenomenon to be clearly recognized was the co-evolution of the eucalypts with the current pattern of bush fires in Australia (Mount 1964). It is, however, unclear whether the *collectively* constructed features of a species' niche can be regarded as part of each *individual* organism, and hence whether they can be assigned biological functions and teleosemantic information content. More theoretical work is needed on the interpretation of this form of inheritance.

4. Stability and Inheritance

Opponents of the parity thesis, such as Maynard-Smith, are well aware of the role of non-genetic factors in development and have a standard strategy for discounting them. This strategy is to question the intergenerational stability of non-genetic factors, and/or their potential to produce the kind of variation upon which natural selection can act. Thus:

> "Differences due to nature are likely to be inherited, whereas those due to nurture are not; evolutionary changes are changes in nature, not nurture; traits that adapt an organism to its environment are likely to be due to nature." (Maynard Smith 2000, 189)

> 'The special status of genetic factors is deserved for one reason only: genetic factors replicate themselves, blemishes and all, but non-genetic factors do not." (Dawkins 1982, 99)

Clearly, the phenomenon of epigenetic inheritance defeats the simplest version of this defense. Developmental systems theory argues that we should define ‘inheritance’ so that something is inherited just if it passes from generation to generation in such a way that evolution can act on its variant forms. Hence, every element of the developmental matrix which is reliably replicated in each generation and which plays a role in the production of the evolved life-cycle of the organism counts as something which is inherited (Griffiths and Gray 1994). The more conservative 'extended replicator view' goes some distance in the same direction (Sterelny, Dickison et al. 1996). It follows that genes cannot be singled out as the sources of developmental information on the grounds that they and they alone persist through lineages long enough for cumulative selection to act upon them. Lineages can be selected for having good methylation patterns or good symbionts or

being imprinted on a good host. These features can persist for evolutionarily significant periods of time. Some other factor needs to be added to the insistence on stable inheritance in order to defend the special, informational status of genes. One candidate is the idea that genetic causes are unique in being 'symbolic' or 'semiotic'. In his most recent discussion, Maynard Smith claims that genes are symbolic and other developmental factors are not, because there is no intrinsic connection between the nature of the gene and the developmental outcome it produces (Maynard Smith 2000, 185). In his commentary on Maynard Smith, Peter Godfrey Smith points out that the apparent difference is generated by looking at the *distal* effects of genetic causes, which depend on an array of other causal factors, while concentrating on the *proximal* effects of non-genetic causes, which, like the proximal effects of DNA, are uniquely determined by physical laws (Godfrey-Smith 2000, 203). In response to this criticism, Maynard Smith introduces a second role for biological teleology in his account of information. A cause is semiotic or symbolic when it has its effect via an 'evolved receptor' that has been selected to confer one of many possible causal 'interpretations' on the signal (Maynard Smith 2000, 215). Whatever its virtues as an analysis of the nature of biological information, this new suggestion fails to generate a principled difference between genetic and epigenetic causes. For example, the mechanisms of habitat and host imprinting described above are clearly 'evolved receptors' in Maynard Smith's sense. They have been selected to confer a highly specific 'interpretation' on an otherwise meaningless chemical input, but in these cases the signal that does not take the form of a DNA sequence.

There is something fundamentally puzzling about Maynard Smith's decision to advocate a teleosemantic interpretation of information talk in biology. If genetic information is teleosemantic information then it can have only a distant relationship to the actual genetic code. The genetic code clearly does not derive its semantic aspect from teleology. The selective history of a sequence of DNA is entirely irrelevant to its meaning in the genetic code. AAA means Lysine even when it is part of a section of junk DNA with no selection history or when it has been inserted by an incompetent biotechnologist who intended it to mean Leucine. This problem cannot be fixed by regarding the biological functions of these individual DNA codons as derived from the general functions of that type of codon, since it is generally supposed that codon to amino acid pairings themselves are 'frozen accidents' and not adaptations. More tentatively, I would suggest that insofar as information talk in developmental biology is related to concepts of biological function, it is to the ahistorical, causal-contribution notion of function, rather than the evolutionary concept of adaptive function. In an important study of functional language in biology, Ronald Amundson and George Lauder have suggested that the causal-contribution notion of function predominates in sciences such as physiology and anatomy (Amundson and Lauder 1994). Like physiologists and anatomists, developmental biologists are primarily concerned with proximal explanations (how mechanisms work at the current time) rather than ultimate explanations (why they evolved). So teleosemantics seems doubly inappropriate as an analysis of information talk in molecular developmental biology.

5. Evolvability and Information

At this point I will leave Maynard Smith's most recent, teleosemantic analysis and turn to what I believe is a more promising approach found in his earlier work. Maynard-Smith and Eörs Szathmáry have argued that the genetic inheritance system and cultural transmission in humans are the only two systems which display what they call 'unlimited heredity' (Maynard Smith and Szathmáry 1995). Only these 'unlimited' inheritance systems, they argue, should really be thought of as 'coding' for outcomes. To make their distinction, Maynard Smith and Szathmary argue that most inheritance systems can only mutate between a limited number of heritable states which can be specified in advance. Habitat imprinting, for example, can only lead organisms to choose different habitats and DNA methylation can only choose whether existing genes will be on or off. The genome and language, however, both have recursive structures. Their basic constituents can be put together in many different combinations and these combinations can be of any length. Hence these inheritance systems have an unlimited number of possible heritable states. At first sight it may not clear why this provides a justification for viewing those systems and no others as transmitting information. One way to grasp the thought behind Maynard Smith and Szathmary's proposal is to see them as providing an objective justification for fixing the source/channel distinction so that genes (or culture) are the information source. The other causes of heritable variation are mere channel conditions because they have relatively few alternative settings. Only in genes (or culture) do we find a signal source with enough possible states to signal the vast range of possibilities that evolution needs. The distinction between limited and unlimited heredity systems is an important one, and may provide insights into one of the key innovations that enabled the diversification of

early life into the vast range of forms we see today. I will argue, however, that it is putting too much weight on the distinction to use it, as Maynard Smith and Szathmáry seem to wish, as a basis to claim that developmental information resides in the genes and that other causal factors provide mere material support to decode this information. A developmental systems interpretation, according to which developmental information exists in the whole matrix of material resources that are need to reconstruct a developmental outcome, is equally consistent with the distinction between limited and unlimited heredity systems and with the insights into the evolutionary process that it provides.

The limited/unlimited distinction seems to support the idea that genes are the signal and the rest of the developmental system is a channel because it suggests that the genes have a vastly greater capacity to 'signal' alternative outcomes. In effect, Maynard Smith and Szathmary propose to partition the total number of developmental outcomes that can be generated by a developmental system between the various inheritance systems that make up that system. The number of outcomes allotted to an inheritance system measures its 'limitedness' and the genetic inheritance system is allotted far more outcomes than any other system. But how is this partition of outcomes to be made? Maynard Smith and Szathmáry assume that the number of permutations of DNA codons, or perhaps of entire genes, is the relevant measure for the genetic inheritance system and that some corresponding measure of the number of permutations of physical parts is the appropriate measure for other inheritance systems. But these are not appropriate measures. For any given inheritance system the range of physical changes that count as evolutionary

changes is restricted to those that can be made use of by the rest of the existing system. The main lesson of the major evolutionary transitions that are the focus of Maynard Smith and Szathmáry's work is that evolution creates entirely new kinds of developmental systems that massively expand the possible interpretations of existing developmental resources, including genes. One way to see this is to consider how many evolutionary possibilities come into existence when a 'limited' heredity system undergoes a change. Consider, for example, the evolutionary possibilities that can be 'signalled' by a base pair substitution in the DNA of a eukaryote cell that cannot be signaled by those substitutions in a prokaryote cell. As Maynard Smith and Szathmary themselves describe, that vast swathe of evolutionary possibilities came into existence through evolutionary change mediated by a limited heredity system. Membranes cannot be constructed without an existing membrane template into which to insert newly synthesized proteins. Hence, major changes to the partitioning of the cell require variation to arise through the membrane heredity system, not through mutations of the DNA. In effect, the measure of 'limitedness' that Maynard Smith and Szathmáry adopt allots to the genetic inheritance system all the outcomes that can be generated by making changes to that system across the full range of possibilities for the other systems while allocating to the other systems only the number of outcomes they could produce given one possible genome. That is why it is not an appropriate measure if the limited/unlimited distinction is to be used to judge the capacity of inheritance systems to 'signal' (cause) evolutionary possibilities. It would be no less (and no more) meaningful to allocate to the genetic inheritance system only the range of outcomes it could generate given one state of the other inheritance systems. The other inheritance systems would receive all the

outcomes that could be generated by permuting the genetic system in the presence of the remaining states of that inheritance system.

A closely related point is that the huge potential of differences in genes, language and perhaps pheremones in social insects) to cause new evolutionary possibilities stems from the fact that these differences 'mean something' to the rest of the developmental system. If the rest of the system surrounding the genes were such that indefinitely many base-pair combinations collapsed into only a few developmental outcomes, then the genetic inheritance system would not be unlimited. It is not hard to imagine cellular machinery with this result - the existing genetic code is substantially redundant in just this way. Hence, the unlimited nature of the genetic inheritance system is more accurately seen as a property of the developmental system as a whole and not of the genome in isolation. The language of separate ‘inheritance systems’ can itself be highly misleading in some biological contexts. Methylation inheritance, for example, is a mechanism of gene regulation, making it odd to describe it as a separate ‘system’ from the genetic ‘system’. From a ‘selfish gene’ viewpoint a gene and the methyl groups attached to it are separate replicators, but no more so than any two genes. From a developmental systems viewpoint genes and methylation patterns are separate developmental resources, but they are elements of one developmental system (Griffiths and Gray 2001).

It has never been part of the developmental systems tradition to deny that nucleic acids and natural languages are distinctive elements of developmental systems. The point of the ‘parity thesis’ is to prevent these empirical differences turning into a kind of scientific

metaphysics, as happens when genes are identified with information (or even 'form') and everything else in development with mere matter. This distracts attention from the many ways in which non-genetic resources sometimes play biological roles more usually associated with genes. It also leads to the empirically inappropriate lumping together of very different non-genetic resources (the 'environment'). From a developmental systems perspective, unlimited heredity is simply another important, empirical property of DNA, and perhaps of other resources, such as pheremones in their role in causing caste-differences in the morphology and behavior of social insects [i].

6. Developmental Switches and Bioinformatics

It is clear that an analysis of 'information' such that genes are the sole or main bearer of developmental information has still not been adequately defended. It is less clear why such an analysis is needed. Developmental systems theorists and other critics of 'information talk' do not deny the unique biological role of genes as protein templates, nor the central role of gene products in development, nor do they deny 'molecular Weismannism' – the prohibition on the inheritance of acquired characteristics via the genetic material (with the exception of a few retro-copying phenomena accepted on all sides). Conversely, advocates of the genetic program do not deny the existence, or even the evolutionary importance, of epigenetic inheritance. There are genuine disagreements among commentators on contemporary developmental biology over the relative promise of more program-like and more dynamicist models of gene regulation, but the lines of cleavage among the disputants do not align with those in the debate over genetic information, as Maynard Smith himself remarks (Maynard Smith 2000, 218). In any case,

that dispute does not turn on whether all the information being processed in a developmental program comes from genetic factors.

If the idea that 'biology is an information technology' is neither obviously true nor a 'forced choice' on which biologists must take a stand if they are to do research, why is the idea so widely accepted? Partly, of course, because of the central role of information in the contemporary scientific world-view (Keller 1995; Oyama 2000a). Another widely acknowledged factor is the misinterpretation of Weismann diagrams (Griesemer and Wimsatt 1989; Sterelny and Griffiths 1999, 64; Sterelny 2000). The now standard diagrammatic representation of evolution shows a causal arrow from genes to phenotypes and a causal arrow from the genes of one generation to those of the next (Maynard Smith 1993: Figure 8). There are no causal arrows that represent the influence of organisms on the reproduction of their genes, nor the many influences that organisms exert on their offspring in addition to reproducing their genes. Both the general cultural enthusiasm for information and the modern Weismann diagram are clearly important in giving the genetic program its air of common sense. But there is another factor closely related to the themes of this paper that has been less widely acknowledged. This factor can be represented by a (very bad) argument:

(1) There is a genetic code

(2) In molecular developmental biology there is talk of signals, switches, master control genes, and so forth.

(3) Therefore, the information flowing in (2) is information in the code of (1)

In this blunt form the argument sounds merely frivolous. But many discussions of molecular biology, especially those for a non-technical audience, insinuate something surprisingly close to it. In his recent paper, Maynard Smith presents something quite close to this argument in a section immediately preceding his analysis of biological information, entitled 'Is the Genome a Developmental Program?'. Here are the first and last sentences of the section and a representative passage from in between:

> There is, I think, no serious objection to speaking of a genetic code, or to asserting that the gene codes for the sequence of amino acids in a protein.
>
> However, an organism is more than a bag of specific proteins. Development requires that different proteins be made at different times, in different places. A revolution is now taking place in our understanding of this process. The picture that is emerging is one of a complex hierarchy of genes regulating the activity of other genes. Today, the notion of genes sending signals to other genes is as central as the notion of a genetic code was forty years ago.
>
> Informational terminology is invading developmental biology, as it earlier invaded molecular biology. In the next section I try to justify this usage. (Maynard Smith 2000,187-9).

While not quite my very bad argument, this series of points is clearly meant to suggest that since the genetic code is 'real science' and not mere metaphor, it is only a matter of

time before other information talk in biology becomes real science too. At best, this is a very weak inductive argument, at worst it is equivocation on the word ‘information’ in several different senses. It is worth pointing out that a number of historians of science have suggested the exact opposite: that the history of information talk in molecular developmental biology is one of consistent retreat from literal to metaphoric in the light of increased understanding of molecular processes (Sarkar 1996; Chadarevian 1998).

The very bad argument (hereafter 'VBA') is one aspect of what Sarkar means when he says that loose information talk 'leads to a misleading picture of possible explanations in molecular biology'. (Sarkar 1996, 187). The information metaphor suggests a 'bottom-up' explanatory strategy in molecular biology, rather than a 'top-down' one. A bottom-up strategy tries to infer the developmental significance of a DNA sequence from the sequence itself, looking at the sequence to determine its product and looking at that product and other gene products to determine how they will interact in development. The alternative, top-down strategy begins by studying the developmental process, works out which gene products are involved and uses the sequence of these products to locate the DNA sequences from which they are made. A realistic picture of work in contemporary developmental biology is that it is at least as much top-down as bottom-up. Note also that the top-down strategy automatically locates and takes account of epigenetic causal factors in development, while the bottom-up strategy finds their involvement in developmental processes an obstacle to upward progress.

I recently came across a rather striking example of the use of the VBA to suggest that molecular developmental biology proceeds by a pure bottom-up strategy. Once again, sequential quotations can be fitted into the framework of the VBA:

> One great discovery was that a set of three rungs contains the information to make one amino acid. So the set of rungs AAA will make one amino acid, while the rungs ACG will make a different amino acid
>
> After *[the human genome project]* we can look for the small section of DNA that holds the instructions to grow a new arm. ... molecular biologists studying the common fruit fly found the section of DNA that controlled the growth of its eyes. They learnt to switch on the growth of eyes...
>
> But even when we have mapped the whole of the human DNA, there's another hurdle to jump before we can start growing arms. The DNA is huge! ... That's why the science of bioinformatics has been invented. (Kruszelnicki 1998)

These quotes are from a leaflet issued to persuade students to enroll in a degree in bioinformatics. Bioinformatics is designed to produce biologically literate graduates trained in computational techniques that will allow them to handle the vast amounts of information *about* genes currently being generated by molecular biology. But deft use of the VBA suggests that bioinformatics deals with developmental information *encoded in*

genes - the blueprint for life, and a much more exciting prospect for school-leavers than applied computing.

7. Conclusion: the Code, the Code and noting but the Code

There really is a genetic code, but beyond that most 'information talk' in biology is a picturesque way to talk about correlation and causation. Such loose information talk is equally applicable to non-genetic developmental factors. Information locutions can be used teleosemantically, but that usage too is equally applicable to other factors. It is also unlikely that this, teleosemantic sense is what is intended by most information talk in molecular developmental biology, which is a science of proximate mechanisms not of ultimate origins. There are numerous important differences between what DNA does in development and the roles played by other causal factors, but these differences do not map onto a distinction between informational and material causation. The present atmosphere, in which information talk is only applied to genes, makes that way of talking highly misleading. I have suggested here that it misleads people about the forms of explanation in molecular biology. I also believe that the asymmetrical use of information talk partly explains the persistence of genetic determinism, but that is an argument for another place (Griffiths Forthcoming).

References

Amundson, R. and G. V. Lauder (1994). "Function without purpose: the uses of causal role function in evolutionary biology." Biology and Philosophy **9 (4)**: 443-470.

Chadarevian, S. d. (1998). "Of worms and programs: *Caenorhabitis Elegans* and the study of development." Studies in History and Philosophy of the Biological and Biomedical Sciences **29**(1): 81-105.

Cosmides, L., J. Tooby, et al. (1992). Introduction: Evolutionary Psychology and Conceptual Integration. The Adapted Mind: Evolutionary Psychology and the Generation of Culture. J. H. Barkow, L. Cosmides and J. Tooby. Oxford, New York, Oxford University Press**:** 3-15.

Dawkins, R. (1982). The Extended Phenotype, Freeman.

Dretske, F. (1981). Knowledge and the Flow of Information. Oxford, Blackwells.

Economist Magazine (1999). Drowning in data. Economist June 26th, London**:** 97-98.

Godfrey-Smith, P. (1989). "Misinformation." Canadian Journal of Philosophy **19 (5)**: 533-550.

Godfrey-Smith, P. (1999). Genes and Codes: Lessons from the Philosophy of Mind? Biology Meets Psychology: Constraints, Conjectures, Connections. V. G. Hardcastle. Cambridge, M.A, MIT Press**:** 305-331.

Godfrey-Smith, P. (1999). The Semantic Gene, Manuscript.

Godfrey-Smith, P. (2000). "Information, Arbitrariness, and Selection: Comments on Maynard Smith." Philosophy of Science **67**(2): 202-207.

Gray, R. D. (1992). Death of the gene: developmental systems strike back. Trees of Life: Essays in the Philosophy of Biology, Griffiths, P (Ed) Kluwer, Dordrecht**:** 165-210.

Gray, R.D (In Press) Selfish genes or developmental systems? Evolution without replicators and vehicles. Thinking about Evolution: Historical, Philosophical and Political Perspectives: Festschrift for Richard Lewontin. R. Singh, K. Krimbas, D. Paul and J. Beatty. Cambridge, Cambridge University Press.

Griesemer, J. R. and W. C. Wimsatt (1989). Picturing Weismannism: A Case Study of Conceptual Evolution. What the Philosophy of Biology Is. M. Ruse. Dordecht, Kluwer Academic Publishers**:** 75-137.

Griffiths, P. E. (Forthcoming). The Fearless Vampire Conservator: Philip Kitcher on Genetic Determinism.

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

Griffiths, P. E. and R. D. Gray (1997). "Replicator II: Judgment Day." Biology and Philosophy **12**(4): 471-492.

Griffiths, P. E. and R. D. Gray (2001). Darwinism and Developmental Systems. Cycles of Contingency: Developmental Systems and Evolution. S. Oyama, P. E. Griffiths and R. D. Gray. Cambridge, Mass., MIT Press**:** xxx-xxx.

Griffiths, P. E. and R. D. Knight (1998). "What is the Developmentalist Challenge?" Philosophy of Science **65**(2): 253-258.

Immelmann, K. (1975). "Ecological significance of imprinting and early learning." Annual Review of Ecology & Systematics **6**: 15-37.

Jablonka, E. and M. J. Lamb (1995). Epigenetic Inheritance and Evolution: The Lamarkian Dimension. Oxford, New York, Tokyo, Oxford University Press.

Jablonka, E. and E. Szathmáry (1995). "The evolution of information storage and heredity." Trends in Ecology and Evolution **10 (5)**: 206-211.

Johnston, T. D. (1987). "The persistence of dichotomies in the study of behavioural development." Developmental Review **7**: 149-182.

Keller, E. F. (1995). Refiguring Life: Metaphors of Twentieth Century Biology. New York, Columbia University Press.

Keller, L. and K. G. Ross (1993). "Phenotypic plasticity and 'cultural transmission' of alternative social organisations in the fire ant *solenopsis invicta*." Behavioural Ecology and Sociobiology **33**: 121-9.

Kitcher, P. (In Press). Battling the undead: How (and how not) to resist genetic determinism. Thinking about Evolution: Historical, Philosophical and Political Perspectives (Festchrifft for Richard Lewontin). R. Singh, K. Krimbas, D. Paul and J. Beatty. Cambridge, Cambridge University Press**:** xxx-xxx.

Kruszelnicki, K. (1998). Why you should study Bioinformatics, University of Sydney Science Faculty publicity material.

Laland, K. N., F. J. Odling-Smee and M.W Feldman. (2001). Niche Construction, Ecological Inheritance, and Cycles of Contingency in Evolution. Cycles of Contingency: Developmental Systems and Evolution. S. Oyama, P. E. Griffiths and R. D. Gray. Cambridge, Mass., MIT Press**:** xxx-xxx.

Maynard Smith, J. (1993). The Theory of Evolution. Cambridge, Cambridge University Press.

Maynard Smith, J. (2000). "The concept of information in biology." Philosophy of Science **67**(2): 177-194.

Maynard Smith, J. (2000). "Reply to Commentaries." Philosophy of Science **67**(2): 214-218.

Maynard Smith, J. and E. Szathmáry (1995). The Major Transitions in Evolution. Oxford, New York, Heidelberg, W.H Freeman.

Millikan, R. G. (1984). Language, Thought & Other Biological Categories, M.I.T. Press.

Moran, N. and P. Baumann (1994). "Phylogenetics of cytoplasmically inherited microorganisms of arthropods." Trends in Ecology and Evolution **9**: 15-20.

Mount, A. B. (1964). "The interdependence of the eucalpyts and forest fires in southern Australia." Australian Forestry **28**: 166-172.

Odling-Smee, F. J. (1988). Niche-constructing phenotypes. The Role of Behavior in Evolution. H. C. Plotkin. Cambridge, Mass., MIT Press**:** 73-132.

Odling-Smee, F. J., K. N. Laland, and M.W Feldman. (1996). "Niche Construction." American Naturalist **147**(4): 641-648.

Oyama, S. (2000a). The Ontogeny of Information 2nd Revised Ed. Durham, North Carolina, Duke University Press.

Oyama, S. (2000b). Evolution's eye: A systems view of the biology-culture divide. Durham, North Carolina, Duke University Press.

Papineau, D. (1987). Reality and Representation. NY, Blackwells.

Radford, T. (1997). Genes say boys will be boys and girls will be sensitive. The Guardian Thursday June 12th, London and Manchester**:** 1.

Sarkar, S. (1996). Biological information: A sceptical look at some central dogmas of molecular biology. The Philosophy and History of Molecular Biology: New Perspectives. S. Sarkar. Dordrecht, Kluwer Academic Publishers. **183:** 187-232.

Shannon, C. E. and W. Weaver (1949). The Mathematical Theory of Communication. Urbana, Ill., Univ. of Illinois Press.

Skuse, D. H. and e. al (1997). "Evidence from Turner's syndrome of an imprinted X-linked locus affecting cognitive function." Nature **387**(12 June): 705-8.

Sterelny, K. (2000). "The "Genetic Program" Program: A Commentary on Maynard Smith on Information in Biology." Philosophy of Science **67**(2): 195-201.

Sterelny, K., M. Dickison and K.C Smith. (1996). "The extended replicator." Biology and Philosophy **11**(3): 377-403.

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

Notes

* A version of this paper was read at the International Society for the History, Philosophy and Social Studies of Biology in Oaxaca, Mexico in June 1999. I am indebted to that audience and to Russell Gray and Robert Olby for comments on the manuscript.

[i] Another slightly different interpretation of the limited/unlimited heredity distinction would be to it to argue that only genes have the capacity to generate fine-grained response to selection and thus to give rise to adaptation. This point could be combined with the teleosemantic approach described above to yield the result that since only genes support genuine adaptation, only genes carry (teleo-) information. As Russell Gray has pointed out, this implicitly assumes that epigenetic inheritance systems must have evolutionary potential separately from one another and that only discrete, as opposed to continuous, variations count as alternative states of an inheritance system. The cards are being stacked against extra genetic inheritance in numerous ways. However:

> Extragenetic changes can also be piecemeal and incremental. Just as natural selection can favor combinations of genes at different loci, so selection might favor combinations of endosymbionts. Quantitative variations in cytoplasmic factors, nest design, and habitat preferences could also all be passed on extragenetically. So although combinations of these factors are not unlimited, they can be quite large enough to allow a fine-grained response to selection (Gray, In Press, xxx-xxx)