Experimental Philosophy of Biology: Notes from the field

Karola Stotz†‡

†Cognitive Science Program, 810 Eigenmann, Indiana University, Bloomington, IN 47408, kstotz@indiana.edu

‡ This material is based upon work supported by the National Science Foundation under Grants #0217567 and #0323496, awarded to the author and Paul Griffiths.

Abstract

I use a recent 'experimental philosophy' study of the concept of the gene conducted by myself and collaborators to discuss the broader epistemological framework within which that research was conducted, and to reflect on the relationship between science, history and philosophy of science, and society.

1. Introduction: Experimental Philosophy of Biology and the 'Biohumanities'

A central tenet underlying the work to be described here is that philosophy and science are not clearly separated activities, but aspects of a single inquiry into nature. They are distinguished primarily by the questions they ask, rather than by any restrictions on their 'proper' methods. The new field of ‘experimental philosophy’ (X-phi) pays tribute to this 'continuity thesis’ by bringing empirical work to bear on philosophical questions. Its practitioners have not lost their identity as philosophers through their employment of methods traditionally associated with the sciences; rather experimental philosophers (‘X-philes’) perform experiments in an attempt to discern facts of relevance to philosophical debates. It is part of the burden of such work to show that some philosophical issues turn on an empirical supposition that may in turn be tested. With respect to studying the changing concept of the gene previous research by the author and her collaborator.s has established that it is possible to operationalize questions about conceptual variation in a survey instrument and that the statistical analysis of these questionnaire data reveals the prevalence of particular gene concepts in different biological fields (Stotz, Griffiths, and Knight 2004). The Representing Genes Project at the center of this paper represents an extension of that earlier work.

The paper reflects on the motivation for undertaking the Representing Genes study, and experimental philosophy of biology in general. X-phi of biology is a potentially
important part of what Paul Griffiths has called the 'biohumanities'. The concept of 'biohumanities' is a vision of the relationship between the humanities (including philosophy of science), biology and society. In this vision, the humanities create knowledge about biology. Both the history of genetics and philosophical work on the concept of the gene of the sort described in this paper enrich our understanding of genetics itself. Contrast this to the vision implicit in the idea of Ethical, Legal and Social Implication ('ELSI') research, in which biologists provide the facts and humanists and social sciences work out their implications for society, or to one traditional vision of history and philosophy of science, in which studies of particular sciences are data for conclusions about the nature of science. In the biohumanities vision, good history and sound philosophical may provide resources for addressing ‘ELSI’ issues, but that is not their primary aim. Likewise, we may learn something about the nature of science from our work, but it is enough if we learn something about genetics, ecology or whatever other bioscience we study.

We can discern four aims of research in the 'biohumanities'. First, Biohumanities understands itself as a critical enterprise. Constructive 'science criticism' (Pigliucci and Kaplan 2006, 8) stands back from the urgencies of actual scientific research to reflect on the strengths and weaknesses of current approaches. Thus, as C. Kenneth Waters has remarked, the aims of conceptual analysis in the philosophy of science include “to articulate scientific concepts in ways that help reveal epistemic virtues and limitations of particular sciences. This means an analysis of the gene concept(s) should help clarify the explanatory power and limitations of gene-based explanations, and should help account for the investigative utility and biases of gene-centered sciences” (Waters 2004, 29).

Secondly, science criticism, without necessarily questioning any specific findings of past research, can simply be the “turning-over of stones that had hitherto held their ground” (Moss 2006, 523). History of science points out the 'roads not taken' in science. Without calling into question the data that was gathered, it points out that others sorts of data might have been gathered, or that the data that was actually gathered could have been

---

1http://paul.representinggenes.org/biohum_home.html
interpreted in different ways. Philosophy of science adds to this enterprise by critically analyzing the chains of reasoning that connect specific scientific findings to claims about the broader significance of those findings. This can lead to changes in interpretation that can potentially motivate biologists to reinterpret earlier scientific findings and to prioritize different questions for future research. Ideally, history and philosophy of biology can articulate alternative visions of biology.

Thirdly, good history and philosophy of biology can contribute to the creation of 'critical science communication' through promoting a critical understanding of scientific claims in the sense just outlined and communicating to a wider audience not merely 'what has been discovered', but something of the complexity of the scientific process and the contestability of its findings. To be useful, critical work of the sort described in the last two paragraphs must be 'bioliterate', something that might be roughly defined as engaging with the science at the same level as practitioners, rather than via popular representations. But the broad 'visions' of science in which it results can be expressed in a non-technical way, and can thus make a major contribution to the public understanding of science.

Fourthly, and most generally, biohumanities is concerned with understanding biology. Although it should by now be clear that I think the biohumanities are of potential value to both biology and society, this is not the only possible justification for biohumanities research. Science is fascinating and important, and it is worth understanding even if understanding it does not make it work any better, just as evolution is worth understanding whether or not doing so contributes to crop improvement or drug development. As in the sciences themselves, it is hard to imagine researchers doing their best work without an intrinsic interest in the material they study.

It may be worth reiterating at this point that professional historians and philosopher do not have a monopoly on any of the activities described in this section. Anyone acquainted with the literature in the history and philosophy of biology knows that biologists have made very substantial contributions to both. But when making these contributions, they
are not doing biology, but history or philosophy, or 'science criticism'.

The following sections of the paper will give more substance to these remarks with the specific example of the Representing Genes project conducted by myself and my collaborators. Section two outlines the project. Section three uses the project to introduce the ideas of scientific concepts as 'tools' for research and the experimental philosopher of science as a 'conceptual ecologist'. In accordance with these two metaphors, the aim of the Representing Genes project was to reveal complementary and competing gene concepts and locate them in different areas of biological research. Our conclusions about gene concepts and their epistemic niches are presented in section four. Finally, in section five, we argue for a new vision of the gene suited to postgenomic biology.

2. Representing Genes

The gene’s conceptual variation is the most salient feature of its 100-year history. Philosophical analyses have attempted to both describe different concepts of the gene in use by different biologists and at different times, and to prescribe better ones (Waters 2004). Many philosophers and historians have analyzed this variation in terms of a tension between two aspects of the gene, aspects that have been characterized in several, not necessarily consistent ways:

- An abstract unit of inheritance versus a concrete entity (Falk 2000)
- A functional or top down versus a structural or bottom up approach to the gene (Gilbert 2000).
- A gene P (for phenotype, predictor or preformation) versus a gene D (for developmental resource)(Moss 2003) a or pop gene (of evolutionary genetics) versus a dev gene (of developmental genetics) (Gilbert 2003).

One general feature that might help to unify these competing characterizations would be to distinguish between a statistical and a mechanistic relationship between the gene and phenotypic traits. Others have approached the variation in the gene concept, not by introducing distinctions, but by unifying the variants under a more general concept, such

---

2 A term we have in fact taken from the joint work of Massimo Pigliucci, a biologist, and Jonathan Kaplan, a philosopher.
as schematic gene which yields different specific genes along a continuum from proximal to distal relationships to gene products (Waters 1994, 2000). In my own work my collaborators and I have come to embrace a version of the dichotomy between an abstract, statistical gene and a concrete, mechanistic gene, but have felt the need to introduce a further distinction between a simple stereotype or ‘consensus gene’ versus a gene which embraces the complex and messy reality of the relationship between genome structure and genome function (Stotz, Bostanci, and Griffiths 2006; Griffiths and Stotz 2006).

The Representing Genes project³ was an attempt to empirically assess the impact of the on-going molecular genomics revolution on concepts of the gene (Stotz and Griffiths 2004; Stotz, Bostanci, and Griffiths 2006). The survey instrument developed for the Representing Genes study was designed to explore several of the issues posed by the existence of alternative gene concepts.

Part one of the questionnaire was inspired by a pressing problem posed by the move into the genome-sequencing era, namely which principles to use in order to functionally annotate DNA sequences. The actual practice of genome annotation inspired us to design a simple, annotation-like task to investigate the criteria that lead biologists to annotate a particular DNA sequence as either one gene with several gene products or several genes with a single functional product. This 'simplified annotation' task used graphical representations and descriptions of real DNA transcription events in eukaryotic genomes which challenge various aspects of the classical molecular gene concept. The cases were chosen to allow pair-wise comparisons highlighting differences in the criteria that may influence biologists’ judgments on this issue during annotation. The examples illustrate the flexibility, variability and complexity of ‘genome expression' that complicate defining what genes are. Since common definitions of the gene are insufficient for making this decision, the simplified annotation task is designed to reveal the implicit criteria biologists draw upon in this judgment.

³ http://representinggenes.org
To give a flavor of the work, we will sketch two of the cases, and the conclusion we drew from them. One of the phenomenon not anticipated by the classical molecular conception of the gene is 'antisense transcription', in which the two strands of the DNA double-helix each contain their own genes, running in opposite directions. Another is 'trans-splicing', in which messenger RNAs transcribed from different parts of the genome are spliced together at a later stage to make a single molecule. This means that a single polypeptide may be derived from several loci in the genome. These phenomena come together in two of our cases which were based on events at the Drosophila locus mod(mdg4). This stretch of DNA contains a fairly conventional-looking gene from which the larger part of a messenger RNA molecule is made by the usual cutting and pasting. The final section of the molecule, however, comes in several variations, which are made from various short sections of DNA located some distance away on the chromosome and trans-spliced to the larger section. Some of these are on the same strand of the DNA double helix as the main part of the molecule. Others are on the opposite strand of the double helix, in which case they run in the opposite direction along the DNA. The mechanisms that make the smaller part of the molecule and join it to the larger part are identical in both cases. We find, however, that while biologists predominantly regard the first case as a single gene, they predominantly regard the second case as two genes cooperating to make a single product. What explains this difference in their responses?

In recognition of the difficulties of the classical molecular gene concept, it has been suggested that working biologists employ a consensus gene concept based on a collection of flexibly applied features of well-established genes. On this proposal, there is something similar to a threshold effect for considering a stretch of DNA to be a gene if it has 'enough' of these features, such as an open reading frame, a consensus core promoter sequence (the TATA box), or the existence of RNA transcripts. The originator of this view, Thomas Fogle, has argued that by combining structural and functional features into a single stereotype, the consensus gene concept hides both the diversity of DNA sequences that can perform the same function and the diverse functions of particular DNA sequences (Fogle 2000). In other words, the consensus gene concept inherently distracts from conceptually problematic cases. It means that when biologists come to
annotate a locus at which something unusual is happening, they look for ways to divide the sequence into several different genes, none of which diverges too far from the stereotype. We suggest that just this has happened in our two cases based on the (mod)mdg4 locus. Although the cases in which the alternative terminal exon is trans-spliced from the same strand are not significantly different in terms of molecular mechanisms from those in which it is trans-spliced from the antisense strand, the idea of a single gene with parts running in opposite directions – 'head to head' – is just too unlike the prototypical gene pictured in textbooks, and so in this case they are seen as separate genes. Whether something is one gene or two is thus as much a psychological as a biological matter.

Another part of the survey instrument set out to investigate whether and when, as Lenny Moss has argued (Moss 2003), investigators either start with the conception of a theoretical gene determinately connected to a particular phenotype, or of a concrete molecular gene with a specific sequence and the template capacity to code for many products, depending on how it is transcribed and how its initial product is later processed. We argued that these different starting points would affect how investigators set out to unravel the complex relationship between genes and other molecular factors with the phenotype. Hence the second task asked subjects to assess the value of different research strategies for investigating complex diseases. For each disease we offered four strategies, designed to run along a continuum from focusing on the statistical relationship between gene and phene to entirely giving up on such a relationship in favor of analyzing content-dependent causal pathways between the two. We looked for differences in which strategies were favored by biologists from different backgrounds, and also at whether the choice of strategies changed between human versus animal disease, and for physiological versus psychological disease.

In this section, I have given a brief description of some experimental philosophy of biology. In the next section I turn to the issue of what philosophers can learn from this kind of work.
3. The Philosopher as a Scientist

One motivation for the Representing Genes study was to transcend the limitations of traditional conceptual analysis. There is a tradeoff between the intimate knowledge of part of the science through interaction with particular scientists and the bias of your interpretation of the whole field. Perhaps as a result, philosophers typically produce competing analyses of scientific concepts, and traditional conceptual analysis too often ends with the 'dull thud of conflicting intuitions'. Such problems have produced increasing interest in bringing a new set of tools to bear. Experimental philosophy of science has the capacity to assess these competing analyses against data and to avoid biases introduced by working with a single subdiscipline or a single school of thought.

Such a philosophy ‘in the trenches’ is also in a privileged position to provide the bridge between philosophy and science. The ‘trench’ of the experimental philosopher does not demarcate the line between the humanities and science in the science war, but the empirical frontline in the fight for real knowledge where philosophy and science unite. At least part of philosophy of science has abandoned the idea that its job is to enforce rigor and precision within science through the fixation of scientific meaning. It has been argued that slippage of meaning was essential to the rapid progress of genetics (Rheinberger 2000). Equally gone is Paul Feyerabend’s conceptual anarchism, in which the history of science is little more than a series of changes in the fashionable topics of scientific discussion (Feyerabend 1975). In place of these two models we have come to appreciate that conceptual change in science is rationally motivated by what scientists are trying to achieve, by their accumulated experience of how to achieve it, and by changes in what they are trying to achieve. Empirical science is a powerhouse of conceptual innovation because scientists use and reuse their terminology in a truly “exuberant” way (Rheinberger 2000). The gene concept is a case in point: despite its ever-changing definition, the gene remains on the laboratory bench after a whole century because it has proved a flexible tool⁴. This only makes sense if we think of concepts as tool of research, as ways of classifying the experience shaped by experimentalists to meet their specific needs. Necessarily these tools get reshaped as the scientists’ needs change.

⁴ But see Moss (2006) with his critique of using the ‘gene’ as placeholder for a full explanation of life.
In the study of conceptual diversification, the history of genetics provides a ‘conceptual phylogeny’ of the gene and the Representing Genes project can be seen as an attempt to determine some of the 'ecological' pressures that have caused the gene concept to diversify into different 'epistemic niches'. The next section presents an attempt of describing the phylogeny and ecology of the former and current use of the gene concepts.

4. 'Conceptual speciation events' and the 'epistemic niche'\textsuperscript{5}

The gene was originally defined in the light of the hybridization techniques available to early geneticists. In the absence of any knowledge about the molecular basis of genetics this early ‘instrumental gene’ was a hypothetical entity, an intervening variable between the phenotypes of the parental generation and the distribution of phenotypes in following generations. As new techniques became available and new questions about the structural nature of the gene pressing, the gene was redefined. However, just as old techniques can survive alongside newer ones, old concepts can remain the best tool for the work for which they were originally designed. For example, when a medical geneticist is seeking the 'genes for' a disorder she is looking for traditional Mendelian genes - sections of chromosome whose pattern of inheritance explains the phenotypic differences observed in patients. Translated into molecular terms these sections may turn out not to be molecular genes. Some abnormalities in human limb development, for example, have been tracked down to mutations in a gene on chromosome 7. But recent research suggests that the gene in which the mutation is located plays no role in the development of these abnormalities (Lettice et al. 2002). Instead, embedded in that gene is a sequence which acts to regulate the use of the gene 'sonic hedgehog', about one million DNA nucleotides away on the same chromosome, which is involved in the relevant aspects of limb development. Nothing is gone wrong in either piece of research. It is simply that the molecular gene concept is not a good tool for some kinds of research. The instrumental, Mendelian gene remains the best tool in fields like medical genetics and population genetics. So while a particular scientific concept reflects the scientific knowledge at a

\textsuperscript{5} For a more detailed explanation of the points in this and the next sections, see (Griffiths and Stotz 2006, in press; Stotz 2006).
point in time, this alone cannot explain the parallel use of several different concepts. For a full understanding of that phenomenon we need to see scientific concepts as tools for research, as much as glassware, microscopes or scales.

In the 1960s molecular biologists believed they had arrived at a single molecular concept of the gene, which united the structural and functional aspects of the gene. The molecular gene is a structure in the DNA whose function is to specify the linear order of elements in a gene product (RNA or polypeptide). This is reflected in the Central Dogma of Molecular Genetics, which claims that the genetic information, the linear sequence of nucleic acid bases, specifies the linear order of the gene product, with no feedback mechanisms allowed (Crick 1958, 1970). In the light of today’s knowledge about the ways in which a limited number of DNA sequences is used to create a vastly greater 'transcriptome' of gene products, the sequences we count to arrive at the claim that there are about 25,000 human or 14,000 Drosophila genes are best regarded as stereotypical genes – sequences that fit a stereotype of how DNA plays the gene-role. The cases that inspire the stereotype are the simple cases of bacterial transcription and translation that were used to derive our basic understanding of molecular genetics in the 1960s. The prototype is undermined by heavily edited mRNA transcripts derived from 'cryptogenes' or by cases of trans-splicing, in which the linear order of the product is no longer mirrored in the linear order of nucleotides in the DNA.

How does a scientist today decide where one gene starts and another stops? Since one gene can code for many different products when expressed in different ways or in cooperation with other sequences, there is no principled answer. The tip of the iceberg of the complexity of gene expression is the common process of alternative splicing. It was the first mechanism detected to seriously undermine the one gene - one polypeptide hypothesis enshrined in the classical molecular gene concept. However, the widespread convention of molecular geneticists is to define alternatively spliced genes as one gene with multiple products. At first surprising, the underlying rationale of the molecular gene

---

66 The majority of genes in higher organisms are alternatively spliced and the current star example of alternative splicing is Dscam (Drosophila Cell Adhesion Molecule), a gene which may produce up to 38,016 different forms of the DSCAM protein.
concept actually explains this convention: A molecular gene is defined as the linear image of a gene product in the DNA (Waters 1994, 2000). Accordingly, if, in the scientist’s judgment, the many different products of a single DNA sequence are sufficiently similar, e.g. protein isoforms with many shared functional subunits, then they are produced from one gene. If, however, they are sufficiently different from one another, e.g. through the process of frame-shifting, then they are the products of two overlapping genes. As we discussed in Section two, a key aim of genome annotation is to find a way of segmenting complex eukaryote genomes into sequences that look reasonably like the prototype of a gene (Fogle 2000).

5. A Postgenomic Gene Concept?
Just as finer work may require more specialized tools, it may be that molecular geneticists are now confronting problems for which the classical molecular gene concept no longer proves useful. For instance, scientists today want to understand how regulated genome expression lead from a ridiculously small number of genes to the explosion of gene products that create and maintain higher organism, especially humans. This will likely require a more modest gene concept in which the structural and functional aspects of the molecular gene are dissociated again. In reality the way in which the DNA contains the image of its product is often akin to the way in which Picasso's cubist paintings contain a fragmented and distorted image of his models. The best trick of the genome, however, the cubists never invented, namely how a single sequence (brushstroke) can be part of many genes (paintings). We now know that complex forms of transcriptional and post-transcriptional processing, at least in eukaryote genomes, are 'business as usual'. Beside the well-known process of alternative splicing, more recently it has been discovered that through trans-splicing coding sequences can be pasted together in a different order, repeated or even pasted in backwards. Thus, as well as getting many products from a single piece of DNA, several pieces of DNA can be used to make a single gene product. These pieces may even be located on different chromosomes. Sequences can be transcribed or translated in different reading frames, or be edited through the insertion, deletion or substitution of nucleotides (see Stotz 2006; forthcoming, for references and description of such cases).
The postgenomic gene concept, rather than covering up any unwanted messiness, welcomes and embraces these complexities in the relationship between DNA and its products as new opportunities to relate a mere 25,000 genes to all the complexities of human development and functioning. Thanks to a large variety of specific complexes of interacting regulatory molecules (cis-acting sequences in the genome, trans-acting factors of gene products, metabolites and other environmental signals) DNA is used in highly time- and tissue-specific ways (Stotz forthcoming). Regulated recruitment and combinatorial control of these regulatory molecules is the mechanism of choice of most organism to control gene expression (Ptashne and Gann 2002). Lenny Moss has described these as 'ad hoc committees' of regulatory molecules whose particular 'membership' reflects the contingent history of the cell up to that time, including the history of the cell's transactions with its environment (Moss 2003). This metaphor is designed to embody the new biology of genome regulation in the same way that the metaphor of a genetic program written in the DNA embodied the biology of the 1960s.

In this vision the ‘gene’ is relieved of its unrealistic and mystical status as the sole embodiment of life with a (unsurprisingly, not very well understood) propensity to ‘work on its own behalf’ (Kauffman 2000; Moss 2006). Instead, genes become prosaic ways to classify the template capacity of certain parts of the genome, a capacity that must be interpreted through a process of gene expression to yield any determinate result. Because of this limited and very context-dependent capacity, the gene is also stripped of its place as the sole unit of inheritance. Predictable expression patterns of parts of the genome are ensured by the reliable reproduction of a developmental niche that regulates the same expression patterns. Inheritance is not embodied in mystical preformations of the phenotype but in the reproduction of the necessary factors of development that will self-organize to reproduce a similar developmental life cycle. Life is not situated in genes but the particular organization of biomolecules that enables the system to maintain itself by reconstituting its own components from the template capacity in the genome, constructing the environmental factors necessary for this to occur, and ultimately reproducing copies of itself.
6. Conclusion
This paper describes empirical/experimental studies in the philosophy of biology in general part of the 'biohumanities'. This field comprises four different but related aims: constructive science criticism, creating alternative visions of biology, critical science communication, and, simply, understanding biology as an object of natural knowledge in its own right. In Section three I outlined how experimental philosophy methods can contribute to this kind of research. Section four and five demonstrate the critical potential of this research when directed at current molecular biology.

References


