Genes*

by PHILIP KITCHER

The gene has been considered to be an undefined unit, a unit-character, a unit factor, a factor, an abstract point on a recombination map, a three-dimensional segment of an anaphase chromosome, a linear segment of an interphase chromosome, a sac of genomeres, a series of linear subgenes, a spherical unit defined by target theory, a dynamic functional quantity of one specific unit, a pseudo-allele, a specific chromosome segment subject to position effect, a rearrangement within a continuous chromosome molecule, a cistron within which fine structure can be demonstrated, and a linear segment of nucleic acid specifying a structural or regulatory product (Carlson [1966], p. 259).

I INTRODUCTION

There is a very natural way to talk about the development of science. As a scientific field progresses, the researchers not only uncover general laws describing the entities with which the field in question is concerned but also arrive at more adequate concepts of those entities. To adapt a famous phrase of Bertrand Russell’s, natural scientists improve their understanding of what they have been talking about, thereby coming to see more clearly why what they have been saying is true. My aim in this paper is to present a general account of conceptual change in science and to illustrate it with one important example. I hope that my discussion will vindicate, at least partially, the commonsense idea that one index of the evolution of a scientific field is its conceptualisation of the entities with which it is concerned. Furthermore, in my treatment of the development of the concept of the gene, I intend to set the stage for the resolution of a hotly debated issue in the philosophy of biology, the question of the relationship between molecular genetics, the theory of Watson, Crick and their successors, and the classical

Received 21 May 1981

* I am grateful to the editor of the British Journal for the Philosophy of Science and to an anonymous referee for their constructive comments on an earlier version of this paper. I would also like to thank Patricia Kitcher for some valuable suggestions.
transmission genetics which grew out of the work of Morgan and his students.

My investigation is thus set in two contexts of recent philosophical discussion. On the one hand, there is a long-standing controversy about conceptual change in science. Two extremely influential writers, T. S. Kuhn and P. K. Feyerabend, have claimed repeatedly that the concepts used by scientists, working in the same field at times separated by large-scale changes in that field, are "incommensurable" (Kuhn [1962], [1977], Feyerabend [1958], [1964], [1965a], [1965b], [1970]). By this, they appear at least sometimes to mean that, at different times, scientists talk about different things—that the "ontology" of the field changes, or that the "constituents of the world" are altered. Critics of Kuhn and Feyerabend have tried to show that the resulting position is incoherent, that the notions of concept and meaning employed by Kuhn and Feyerabend are confused or misguided (Shapere [1966], Scheffler [1967], Kordig [1971], Achinstein [1968], Davidson [1974].) More recently, there have been attempts to apply new insights in the philosophy of language to oppose the Kuhn–Feyerabend thesis. I shall try to show that the approach to meaning and reference offered by Kripke, Donnellan and Putnam (Kripke [1980], Putnam [1975], Donnellan [1972], [1974]), enables us not only to recognise what is wrong with the more extreme claims about the incommensurability of scientific concepts but also to elaborate the natural idea that scientific concepts change and that later theories refine the concepts of earlier theories. Here I shall present in much more detail a view of conceptual change which I offered in my [1978].

The second context is that of the reductionism debate in the philosophy of biology. Since 1953, the field of genetics has been transformed. (See Judson [1979] for an entertaining and lucid review of the main developments.) Molecular biology has clearly done something important for classical genetics. But what exactly is the relation between the new molecular genetics and the classical genetics, articulated by Morgan, Muller, Sturtevant, Bridges, Beadle, Tatum, McClintock, Stadler and a host of other luminaries? The standard philosophical answer to this question has been that classical genetics reduces to (or has been reduced to) molecular biology (Schaffner [1967], [1969], Ruse [1971], Goosens [1978]), but, as a series of criticisms (and concessions) have demonstrated, this solution is highly problematic (Hull [1972], [1974], [forthcoming], Schaffner [1974], Wimsatt [1976], Kimbrough [1979]). Reductionists have attempted to show how genetics fits the standard model of theory reduction advanced in Nagel [1961], or some liberalised version of that model; antireductionists contend that the case of genetics will not fit any nontrivial reduction model, usually maintaining that appropriate "bridge principles" cannot be found. The antireductionists have shown (to my mind, at least) that the reductionist approach will not do. However, that should not lead us to dismiss the philosophical question. We should continue to wonder about the relation
between molecular genetics and classical genetics and to look for a method of
describing clearly what the last three decades of research in the field of
genetics have accomplished.

In this paper I shall try to explain one part of the relationship between
molecular genetics and classical genetics. On the view I shall recommend,
molecular genetics refines the concept of the gene (although we shall see that
it does so in a slightly surprising way). This is only part of the story. I would
also want to argue that molecular genetics deepens the explanations
provided by classical genetics, but to elaborate that claim would require me
to offer an account of theory and explanation in genetics which I am
currently only able to sketch. For the present, I shall be content to uphold
the view that my theory of conceptual change in science allows us to
understand some aspects of the relationship between molecular and classical
 genetics, and that it therefore takes us beyond the debate about reduction
towards a solution of the philosophical problem from which that debate
sprang.

2 REFERENCE AND REFERENCE POTENTIAL

Talk of conceptual change leads to talk of concepts, so that, almost from the
start, those who hope to discuss conceptual change in science appear to be
enmeshed in references to mysterious intensional entities. I believe that we
can manage without such entities. The semantical approach to conceptual
change which I favour is extensionalist: we can adequately describe the
phenomena of conceptual change by charting the shifts in referential
relations between words and the world. One kind of conceptual change is
quite straightforward. Sometimes expressions used by scientists may come
to lose their old referents and to acquire new ones. So, for example, we are
inclined to say that the extension of 'planet' (as used by pre-Copernican
astronomers) did not contain the Earth, but after the acceptance of
Copernican theory, the Earth was included in the extension of 'planet'. Our
inclination is at least approximately right. 'Planet' is one of the most striking
eamples of a term which undergoes referential change in the course of the
history of science. However, we would be wrong to suppose that this is the
only type of conceptual change which can occur, or even that it is the most
important.

A second type of conceptual change does not involve the replacement of
the old referent of a term with a new referent, but rather an alteration in the
mode of reference of the term. Consider a typical scenario. Scientists
introduce into their language an expression which refers to a particular
entity, without being able to provide, in antecedently available language, a
description which would pick out the referent. Later, theoretical develop­
ments enable their successors to specify the entity in question. This kind of
change has occurred with respect to a host of scientific terms: prominent
examples are 'magnet', 'temperature', 'acid', 'compound', 'species'—and,
as we shall see, 'gene'. The reason for describing it as an alteration in the mode of reference is that the acquisition of an identifying description paves the way for a different means of fixing the reference of the old expression. Before they are able to provide an identifying description, members of the scientific community refer to the entity in question by virtue of causal links between that entity and their tokens (I shall have more to say about this shortly). After the description is available, they can fix the referent of a token of the expression as that entity which satisfies the description. So, for example, the reference of 'acid' was originally fixed through events in which paradigm substances were labelled as acids. Now that we have identifying descriptions, the referents of some of our tokens of 'acid' are determined through those descriptions: thus we pick out acids as compounds which are capable of furnishing hydrogen ions.

Obviously, both the kinds of changes which I've discussed so far can occur in tandem. Having used a predicate to refer to a particular set, one may acquire a description which is taken initially to state the condition of membership in that set. It may then transpire that the set determined by the description is not the old referent, and that the old referent included elements which, from the perspective of the later theory, are quite disparate. So one comes to employ the old expression to refer to the set determined by the new description, with the result that there is not only change in mode of reference but the substitution of a different referent as well.

Although the account offered so far captures many of the most striking cases of conceptual change in science, I think that it is incomplete. I have been discussing special cases of a more general phenomenon. Scientific expressions are associated with a complex apparatus which I shall call their reference potential. The reference potential of a term for a community is a compendium of the ways in which the referents of tokens of the term are fixed for members of the community. More exactly, I suppose that members of the community share a set of dispositions to verbal behaviour, containing dispositions to refer in various ways. In general, conceptual change is change in reference potential. The dispositions of the community to use tokens of a particular term to refer in a particular range of ways may vary with time, and this is the phenomenon of conceptual change which has interested historians and philosophers of science. The most dramatic examples are those cases in which the community becomes disposed to use tokens of an old term to pick out a new referent, and those cases in which it acquires a disposition to fix tokens of an old term through a description, where no such referential specification had been possible before. These are just the cases singled out above.

My approach to conceptual change presupposes the thesis that different tokens of the same expression often refer differently. I have argued for this thesis elsewhere (Kitcher [1978]), and I shall review below those points of the argument which are salient to discussion of conceptual change in genetics. Before I do so, it will be helpful to explore in more detail the
notions of “fixing the reference” and of “ways of fixing the reference”, which I’ve employed rather casually in the foregoing.

In recent years, semantic theories which suppose that the referent of a speaker’s term is that entity satisfying a description which the speaker would provide have been severely criticised (Kripke [1980], Donnellan [1972], Putnam [1973]). Many writers have pointed out that it is possible for someone to use a term to refer without being able to produce any (non-question-begging) description, and that, even if a speaker is able to produce an appropriate identifying description, the referent of the term (or of her token of the term) need not be the entity satisfying the description. I can use ‘weasel’ to refer to the set of weasels, even though I can’t provide a description which would distinguish weasels from stoats (and even though I couldn’t tell the weasels from the stoats if I were shown a group consisting of animals of both kinds). Moreover, even if I were to believe that tigers were herbivorous, spotted canines (producing some such erroneous description when asked to identify tigers), it is still possible that I should use ‘tiger’ to refer to the set of tigers.

These criticisms bring into prominence the social character of much of our use of language. I can refer to weasels, and my misinformed counterpart can refer to tigers, because I (and my counterpart) belong to a linguistic community, within which there are people, experts, who can distinguish weasels from stoats and can identify the tigers among the quadrupeds. My references are parasitic on those of the experts. Yet, even in the case of the experts, we are not required to suppose that they can provide descriptions which identify the referents of the terms. It is enough for them to be able to discriminate, by whatever means, the entities to which the terms apply—or even to decide, by fiat, that the term shall apply to a paradigm object and anything “sufficiently similar” to it. Thus, in place of the picture of a speaker having in mind a description determining the referent of the expression he employs, we are offered a rival picture. The speaker is viewed as a member of a community of speakers, so that the referents of his terms may be determined not by any description he has in mind but as the referents of the terms of those from whom he acquired the terms, and ultimately as the referents of the terms of the experts or those who introduced the terms. Moreover, even the references of experts or term-introducers may be accomplished without the mediation of descriptions, either because of powers to discriminate or because of direct attachment of words to objects, as, for example, in baptismal ceremonies.

This new picture isolates important elements of language use. However, it should not lead us to disregard the intentional components which were central to the traditional account. Indeed, unless we give some role to the intentions of language users, we shall fall prey to obvious difficulties. Consider, first, the possibility of introducing a term for a biological kind, such as ‘tiger’, along the lines I have indicated. We imagine some courageous spirit ostending a tiger and attaching the name to the kind
exemplified by the organism before him. Trouble threatens our account because no single kind is instantiated in that organism. (The trouble is noted in Dupré [1981], p. 76.) What makes the newly introduced expression a name referring to the set of tigers rather than the set of quadrupeds, carnivores, mammals or vertebrates? Once the question is posed, an answer is not hard to find. The speaker “has in mind” a similarity relation which would fix the reference to the kind tiger. More exactly, given the intentions with which the term was introduced, the speaker would be disposed to withhold ‘tiger’ from many quadrupeds, carnivores, mammals and vertebrates. The referent of the new term is the kind that best fits the dispositions to verbal behavior.

Another worry is that, when we abandon the idea that the speaker’s referent is determined by a description which hovers before her mind as she speaks, reference is something which simply happens to a speaker. Is it possible for someone simply to blurt out a sound and fortuitously refer to some unenvisaged object? To allay the concern it suffices to note that speakers’ intentions are to have a role in the determination of reference. If someone produces sounds with the general intention to agree in usage with her fellows, then that is enough to secure a connection between the sounds emitted and a referent. Where such general intentions are lacking, the speaker functions more like a parrot: however much sound or fury there may be, nothing is signified.

The new picture of reference does not ignore the role of speakers’ intentions. It focuses on intentions which were not previously recognised or, if recognized, were not credited with their proper role. For most of our linguistic performances, the general intention to conform to the usage of others is far more important than any intention we may have to refer to whatever fits a description which we would be inclined to produce. That is not to deny that a speaker’s behaviour may sometimes make it clear that she does not intend to follow common usage. For example, she may say ‘I don’t care how others use the term; this is how I intend to use it’. Under such circumstances we are prepared to take her as referring to an object which is not the common referent but which fits the description she gives.

When we are examining scientific usage, the intention to conform is by no means the only one that has to be taken into account. Scientists usually also have the general intention of referring to natural kinds, and, in recognising this intention, we sometimes construe the descriptions which they offer as mistaken identifications of the referent rather than as successful specifications of a different referent. However, as I’ve already noted, there are occasions on which scientists are concerned to provide an explicit specification to explain what they are talking about, occasions on which they intend that the referent should be whatever satisfies a particular description or whatever was causally linked to the production of a token on some particular earlier occasion. We can view the scientist as intending to obey three maxims:
Conformity: Refer to that to which others refer.
Naturalism: Refer to natural kinds.
Clarity: Refer to that which you can specify.

Obviously, there are many circumstances under which these maxims conflict, so that the scientist has to “choose” among them. (My use of inverted commas here acknowledges the point that there is no process of conscious choice. Indeed, the intentions under discussion here are frequently not conscious.) Because the choice can be made in different ways of different occasions of utterance, different tokens of the same type can easily refer differently.

We are now in a position to integrate the insights of the new picture of reference with an idea which was fundamental to the older picture. A theory of reference should help us with any occasions of utterance. The speaker produces a token. The theory of reference should supply a criterion which, given sufficient nonsemantic information, can be used to determine to what, if anything, he referred. The theory of reference which I favour claims that the correct hypothesis about the reference of a speaker’s token is that hypothesis which best explains why he said what he did. Successful ascriptions of reference accord with a principle which Richard Grandy has aptly dubbed the ‘principle of humanity’ (Grandy [1973]). This principle enjoins us to impute to the speaker whose references we are trying to understand a ‘pattern of relations among beliefs, desires and the world [which is] as similar to ours as possible’ (ibid. p. 443). Here we must recognise that the particular intentions of the speaker on the occasion of utterance play an important role. The best explanation of the utterance tells the correct story about the speaker’s intentions in making the utterance, relating those intentions to the external circumstances of the utterance and to the speaker’s verbal behaviour. A talented historian does this, perhaps without ever formulating explicitly what is being done. The task is to use our understanding of the pattern of relations among mental states and the world which is common to humanity, together with the available data about the speaker’s (writer’s) environment and behaviour, to identify the intentions which were operative on the occasion of utterance and thus to construct an explanation of the production of the tokens produced. The mistake of the old picture of reference is that it gave undue prominence to one particular feature of the occasion of utterance. It was assumed that there would always be some description which a speaker would associate with her expression, and that the disposition to produce this description was of paramount importance in identifying the referent. However, I think it would be equally mistaken to suppose that some other feature of the speaker’s psychology—say the intention to obey Conformity—always has precedence. Especially when we are dealing with the utterances of those who are attempting to develop science in new ways, we should recognise the possibility of a break with traditional usage. In some cases, the overriding intention is to refer to a
natural kind including a particular object or substance, and, in such cases, the descriptions which may be introduced to characterise the referent should not be regarded as determining it. (However, certain partially identifying descriptions which the speaker is inclined to give may indicate which kind of the many exemplified in the ostended object is to be the referent.) In other cases, in the interest of ensuring that the intended reference is clearly understood, a scientist’s dominant intention may be to refer to whatever satisfies a particular description, even though it may turn out, unluckily, that this is not a natural kind. 1 Traditional theories of reference have tended to assimilate all cases of scientific usage to the latter type. In rejecting the assimilation, we do not need to make the equally unwarranted assumption that scientific utterances are always governed by the intention to obey either Conformity or Naturalism.

It is important to recognise that the intention to satisfy Clarity rather than Conformity or Naturalism is sometimes eminently rational. One goal for the scientist is the development of a language in which he will not only refer to natural kinds but also be able to connect his expressions by providing many alternative ways of specifying the referents of his terms. In the pursuit of this goal, it is not always appropriate to introduce expressions referring to genuine kinds without any ideas about the membership conditions of those kinds. Naturalism is legitimately sacrificed to Clarity when the scientist wants to establish a particular usage for the purpose of obtaining explicitly limited results. Thus, at the early stages of research in a field, ideas about the genuine kinds may be so uncertain and indefinite, that progress is made more quickly by introducing expressions whose reference is fixed descriptively. Operational definitions, construed as descriptive specifications of the referents of tokens of terms, can be a crucial first step in the assembly of data that will eventually help in building more adequate concepts. Although the operationalist programme made excessive claims, the view of reference proposed here allows us to credit it with an important insight. By doing so, it permits us to justify the claims of the many practising scientists in immature fields who find philosophical critiques of operationalism intellectually compelling but methodologically irrelevant.

We can now give a clearer description of the idea of reference-fixing. I suggest that the referent of a speaker’s token is that entity which figures in the appropriate way in a correct explanation of the production of the token. To cash out the notion of figuring in an appropriate way, we should consider the various forms that the explanation may take. Any such explanation will

---

1 We can dramatise the point by imagining two kinds of soliloquy in which the scientist may engage. The first type consists in saying: ‘I intend to pick out a natural kind, and I hope that this description characterizes a kind; if it turns out that it doesn’t, I’ll abandon the description.’ The second consists in saying: ‘I intend to pick out whatever satisfies this description, and I hope it’s a natural kind; if it turns out that it isn’t, I’ll stick with the description.’ My point is that there are some utterances which are best understood by attributing the first type of soliloquy, and some which are best understood by attributing the second.
consist in the description of a sequence of events whose terminal member is the production of the token. **Direct** explanations will relate this terminal event to a governing present intention to refer to whatever satisfies a particular description or to a governing present intention to refer to an entity with which the speaker is currently in direct causal contact (or to a natural kind containing such an entity). **Indirect** explanations will relate the production of the token to a present intention to conform either to the usage of others or to the prior usage of the speaker; such explanations will thus link the terminal event to previous occasions of production of tokens, and we shall repeat the procedure until we arrive at an utterance for which a direct explanation can be given. In all cases, the reference of a speaker’s token is fixed through an **initiating event**, that is an event in which someone (not necessarily the speaker whose reference is under consideration) is either in causal contact with an entity or intends to refer to whatever satisfies a description. We can replace the notion of ways in which the reference of tokens of a type can be fixed with the idea of a collection of initiating events for productions of such tokens. Hence I shall regard the reference potential of a term-type for a speaker as the class of events which, given the speech dispositions of the speaker, can initiate productions of tokens of the type.

Many champions of the new picture of reference seem to suppose that there is one event which initiates all subsequent productions of tokens of a given type. They write as though a word is originally linked to the world in a baptismal ceremony and all later uses of the word achieve their reference through the original ceremony. I suggest a rival picture in which the connections of terms to the world are frequently renewed and extended. It is now possible for me to explain why scientific terms frequently have heterogeneous reference potentials, that is, reference potentials which include diverse initiating events. Scientists who engage in different projects frequently find it useful to initiate their tokens by different events. For the purposes of one kind of research, it is convenient to identify the referent of a token as that which meets a particular description; for other purposes, a different description or a different kind of causal interaction may be more appropriate. An obvious illustration of this is the use of terms for chemical substances. A substance can often be generated in several different ways, and workers in one subfield may find it useful to fix the referent of its name through one mode of production, while workers in another subfield employ another mode of production or an identifying description. When this kind of thing occurs (as it very frequently does) it is reasonable to claim that the term-type in question is **theory-laden**. In using tokens of the same type to refer via different kinds of initiating events, the community presupposes that the entities singled out (by description or causal connection) in these events are the same. (In my [1978], I adapt a classic argument of Hempel’s

---

1 As the referee has pointed out to me, the emphasis on baptismal ceremonies may be quite appropriate in discussing some scientific terms, such as the species names introduced by taxonomists.
346 Philip Kitcher

([1966] pp. 91–7), to show that the risk always has to be taken. Thus I endorse the general thesis that scientific terms are theory-laden.) Of course, I don’t mean to suggest that the presupposition needs to be explicitly formulated by the community.

The thesis that scientific term-types have heterogeneous reference potentials is the key to solving the problem about incommensurability mentioned in I. If by “incommensurability” Kuhn and Feyerabend simply mean mismatch between the reference potentials of expressions in the languages of different theories, then they are quite correct to hail the examples they favour as cases of incommensurability. However, this type of incommensurability does not involve any inability of the language of later theory to specify the referents of the tokens produced by earlier theorists.¹

It is easy to see how the development of science can cause changes in the reference potential of an expression. Discoveries and acceptance of new hypotheses can enlarge the reference potential by suggesting new kinds of causal interaction or new descriptions through which the reference of tokens of the term can be fixed. Alternatively, as in the examples of the revolutions discussed by Kuhn and Feyerabend, scientists may learn that what they had believed to be causal interactions with the same entity are actually encounters with different entities, or that descriptions taken to be coextensive are not coextensive. Such experiences can lead to contraction of the reference potential. When we put both types of change together, we see how radical conceptual change is possible as the result of a continuous process. At a later time, the reference potential of a term may share no common element with the reference potential at an earlier time. Yet the evolution of the concept may be continuous, in that, at intermediate times, the reference potential of the term may contain elements from both the nonoverlapping classes.

So far, I have been trying to explain how we might give sense to the idea that different tokens of the same type refer in different ways and how this idea may be used in understanding conceptual change in science. Yet, throughout my discussion, I have talked about a community of speakers. Before closing this presentation of my theory, it is important to consider how such linguistic communities might be identified.²

With respect to a particular expression type, two speakers belong to the same linguistic community if they are disposed to count exactly the same events as initiating events for production of tokens of the type. This agreement can come about in many different ways. At one extreme, the speakers might both be fully aware of all the events they would count as initiating events for the type, so that their agreement could be established in principle by having them describe the relevant events. At the other extreme,

¹ I now think that my [1978] may have been a little uncharitable to Kuhn and Feyerabend. Their papers seem to oscillate between two notions of incommensurability, one which my account reveals as inadequate and the innocuous idea of mismatch of reference potentials.

² Here I am indebted to the referee.
they might have no clear conception of how their references are fixed, and simply rely on some third party, so that they would concur in virtue of a general disposition to follow the references of this expert. Plainly there are intermediate cases.

When we study the utterances of scientists of a particular time, all of whom use the same expression, we may find a number of linguistic communities. These communities may be related, in that one group may adopt a slightly more inclusive reference potential for the term than another. Once this point is recognised, we can refine further our theory of conceptual change. I have pretended that there is one linguistic community, and that the study of conceptual change charts the evolution of its reference potentials. Now I suggest that the modification of scientific concepts may involve the alteration of the reference potentials of several linguistic communities, possibly with cases of fusion and fission along the way.

I want to emphasise two points. First, two members of a linguistic community may differ greatly in their beliefs. What is crucial is that they agree on the ways in which the referent of a term should be fixed. Second, not all community shared beliefs which use a particular term may be employed in fixing the reference of that term. It is quite possible that each member of a linguistic community should be prepared to assert that the things referred to by a particular term lack a particular property, and yet use that term to refer to entities which have that property. So long as the belief is not used to fix reference, a false belief may prevail throughout a community. In both of these envisaged cases I rely on a distinction between beliefs which are employed in reference-fixing and beliefs which are not. This distinction is generated by the general theory of reference outlined above, which provides a criterion for determining the initiating event for a speaker’s token.

Here, finally, is a synopsis of my view. Conceptual change in science is to be understood as change in reference potential. The reference potential of a term for a speaker is the set of events which a speaker is disposed to admit as initiating events for tokens of that term. A linguistic community, with respect to a term, is a set of individuals disposed to admit the same initiating events for tokens of the term. An event is the initiating event for a token if the hypothesis that the speaker referred to the entity singled out in that event provides the best explanation for her saying what she did. Explanations are judged by their ability to provide a picture of the speaker’s intentions which fits with her environment and history and with the general constraint of the Principle of Humanity. Three kinds of intentions are prominent: the intention to conform to the usage of others, the intention to refer to natural kinds, and the intention to refer to what can be specified.

3 GENES, CHROMOSOMES AND FUNCTIONAL UNITS

I intend to use the machinery developed above to show how we can advance beyond the debate about reductionism in genetics. I shall examine the ways
in which the reference potential of 'gene' has changed in response to experimental and theoretical discoveries and innovations, specifically advances in molecular biology. This will provide us with a fine-grained analysis of what molecular biology has done for the concept of the gene.¹

From the very beginning, geneticists were talking about chromosomal segments. Whether they used the expression 'unit character', 'factor', 'gene' or whatever, the pioneers of classical genetics referred to sets whose members are pieces of chromosomes. It does not matter that some early geneticists were agnostic about the physical basis of "factors" or opposed to the chromosome theory. Their utterances are best explained by taking them to refer to sets including the entities whose effects they followed in their breeding experiments. Those entities were chromosomal segments. My theory provides a straightforward account of their references.²

However, though we must grant that, in one sense, geneticists have always been talking about the same things, namely chromosomal segments, we must also recognize that different tokens of 'gene' have referred differently. On different occasions of usage, different sets of chromosomal segments have been picked out. The major question to be asked in charting changes in the concept of a gene is 'What is the principle of segmentation which determines the referent of this token?'

An appropriate place to begin a discussion of the development of the concept of gene is with the list of approaches to the concept which I quoted at the beginning of this paper. I am not going to try to decide the status of all the descriptions Elof Carlson gives: that would be an enormous, and not terribly profitable task. Instead I am going to pick out two central types of approach, both of which involve a partially functional characterisation of the gene. One of these picks out genes by their function in producing macroscopic effects, or, at least, phenotypic effects. The other identifies genes by focusing on their immediate action. I shall argue that the former approach, prominent in classical genetics, gives rise to a number of different gene concepts and that the use of these concepts brings certain problems for the science. The latter approach, which receives a precise formulation in molecular biology, shows how to solve the difficulties emerging from the

¹ The account I shall offer depends on the following sources: Strickberger [1976], Whitehouse [1965], Watson [1976], Peters [1959], Carlson [1966], Muller [1962], Morgan et al. [1915], Dunn [1965], Sturtevant [1965]. Strickberger provides a good account of the basic genetics, Watson's text is a lucid survey of molecular biology, and Whitehouse links his presentation of genetic theory to the historical development of the subject. My chief debt is to Carlson's magnificent 'critical history' ([1966]). Many of the points I shall make could be elaborated further with examples from his book.

² Given my definition of 'linguistic community', those who opposed the chromosome theory belonged to a different linguistic community from those who used the identification of genes with chromosomal segments to fix the reference of 'gene'. Nevertheless, members of the former group referred to chromosomal segments because, despite their belief that genes are not chromosomal segments, the best explanation of why they said what they did is to take them to refer to such segments. What sets them apart from the "orthodox" community is their refusal to admit a particular way of referring to chromosomal segments, namely by using 'chromosomal segment' as part of a reference-fixing description for 'gene'.
classical concept and how and to what extent classical usage can be unproblematically preserved.

According to many philosophical accounts, classical genetics considered genes in three different ways. (Schaffner [1967], Field [1973].) Genes were taken to be the units of function, the units of recombination and the units of mutation. Benzer’s work on genetic fine structure then showed that the first characterisation is equivalent neither to the second nor to the third. Functional units, cistrons, are typically far larger than units of recombination, recons, or units of mutation, mutons. (Chromosomes which pair at meiosis can exchange material. Recombination is the process in which these exchanges occur. Units of recombination are chromosomal segments which, if exchanged at all, are exchanged as wholes. Mutation, by contrast, is a structural change within an individual chromosomal segment, independent of its interaction with another chromosome. Units of mutation are the smallest segments within which such changes occur.) The philosophical accounts tell a simple story, but the simplicity is deceptive. In the first place, the classical concept of a gene involves the use of recombinational and mutational criteria to specify what is meant by a ‘unit of function’. Secondly, although some classical geneticists believed that recombination cannot occur within genes, it needs to be argued that the reference of ‘gene’ is fixed by some such description as ‘a chromosomal segment within which recombination doesn’t occur’. Thirdly, the belief that genes were units of mutation was definitely a minority view among classical geneticists. Most geneticists believed that mutation doesn’t involve a change in the entire gene. (An exception seems to have been the physicist-geneticist Delbrück who conceived of genes as molecules capable of existing in a number of states, viewing mutation as a change from the normal molecular state.) A natural interpretation of cases of multiple alleles is that these consist of changes at different places in the wild-type gene, and this interpretation was taken seriously by geneticists from 1915 on. Moreover, Muller’s studies of mutation, which are far more extensive than those undertaken by anyone else, claim repeatedly that mutation doesn’t involve change of the entire gene. (See Peters [1959] pp. 109, 153; Carlson [1966] pp. 75, 86, 102-3, 133–8.)

Mutational and recombinational criteria became part of a principle for segmenting chromosomes quite early in the history of classical genetics. Mendel and those who rediscovered his work thought of genes (factors, unit-characters) as functional units. The simplest characterisation of genes

1 Ironically, in his [1978], Goosens, who correctly recognises that classical genetics rejects the thesis that mutation involves a change in the whole gene, claims that it would have been unreasonable for classical geneticists to suppose that recombination can’t occur within genes. Classical geneticists did hold that intergenic bonds were weaker than intragenic bonds, and there was no obvious reason for dismissing this idea. Moreover, modern studies of the mechanism of recombination have indicated that the idea even contains a grain of truth. Although we now recognise the possibility of intragenic recombination, it appears that the enzymes that “cut” the chromosomes and “splice” them initiate crossing-over at the ends of cistrons. (See Whitehouse [1965] pp. 367–9.)
along these lines would be to take a gene to be an entity which produces a phenotypic trait. There are two important areas of vagueness in this specification. What exactly is meant by production? What counts as a phenotypic trait? Waiving these questions for the moment, we should recognise that so simple a principle of segmentation does not accord with those early geneticists (including Mendel himself) who allowed for the existence of 'compound characters', phenotypic traits affected by more than one pair of genes. Consider, for example, the phenotypic characteristic of eye colour in *Drosophila*. Morgan and his successors fashioned a concept of the gene which regards this characteristic as dependent on a large number of genes located on different chromosomes. Adherence to the simple functional criterion would have promoted a concept of gene according to which the 'eye color gene' would be a complex aggregate of scattered segments.

In its early usage, 'gene' (or 'factor') referred to a set of chromosomal segments each of which plays a functional role in the determination of a phenotypic trait. To specify this functional role is a tricky matter. Let's begin with the idea of a *gene complex*. A gene complex is an aggregate of that chromosomal material whose nature determines the form taken by some phenotypic character. A *division* of a gene complex is a set of chromosomal segments whose sum is the gene complex. A division is *optimal* with respect to a set of data obtained from breeding experiments if and only if it is possible to construct enough genotypes from the elements of the division to equal or exceed the number of distinct phenotypes observed, and there is no coarser division which will do this. Finally, a *gene* is an element of a division of a gene complex which is optimal with respect to an ideal set of data.

We can investigate the notion of an optimal division more thoroughly by introducing a key feature of the early procedures for identifying genes, the use of mutant characters. A *wild-type* gene complex is an aggregate of chromosomal material which produces a normal phenotype. A *mutation site* is a connected region of chromosomal material such that change in that region of a wild-type gene complex produces a specific deviation from normal phenotype. An optimal division divides the gene complex into segments each of which contains one (or more) mutation sites, but it does not divide those mutation sites which are not separated in the experimental data. All of this is a formal version of the traditional procedure, systematically used by Morgan and his successors, for cashing out the intuitive idea that a gene is a chromosomal segment which plays a distinctive role in the production of a phenotypic trait.

The history just presented, coupled with the theory of reference of 2, enables us to understand some of the references of Morgan, Sturtevant, Bridges and Muller, provided that we can resolve some preliminary questions. Are we to take the class of phenotypic traits to include characteristics which we can now detect, or should we restrict the class to comprise just the kinds of traits investigated by the *Drosophila* group? In
determining mutation sites, should we appeal to our own methods of detecting mutants or to those methods available to Morgan et al.? Finally, in assigning mutation sites to genes, should we use the breeding experiments which were actually performed or those which could have been performed? It is quite possible that recombination between close mutation sites would not occur in the experiments run by the early Drosophila geneticists, because too small a sample of organisms was used. Under these circumstances, the theorists in question would be led to assign to the same unit sites between which recombination can occur. We can respond in one of two ways: either their use of the actual experimental data is given primacy, so that the reference of ‘gene’ is taken to include a segment covering both sites, or we can favour the theoretical specification of mutational sites as belonging to distinct genes if recombination can occur between them, thereby viewing the theorists in question as making a mistake about genes.

It should come as no surprise that I regard the references of the Drosophila group as context-dependent. ‘Gene’ quickly acquired a heterogeneous reference potential. All the questions raised in the last paragraph should be resolved in different ways in different contexts. The point can be made clearly by considering the final issue. In 1913, Sturtevant concluded that the genes for white and eosin eyes in Drosophila were completely linked and that white and eosin were thus both alleles of the wild-type gene. If experiments with a sufficiently large sample of flies had been conducted, then crossovers between the alleged alleles would probably have been discovered (Carlson [1966] p. 64). I suggest that, in making the original claim that white and eosin are alleles Sturtevant was using ‘gene’ to refer to segments within which crossing over cannot occur, so that the claim was mistaken. However, the mistaken claim set the stage for subsequent usage. When we later find geneticists using the white-eosin allelism as a basis for arguments about the possibility of different variations within the same gene (to cite just one example of a type of utterance in which the deliverances of the actual experiments are taken as primary) their utterances should be understood by regarding ‘gene’ as referring to a set which includes a segment covering the white and eosin mutation sites. Sturtevant’s mistake enlarges the reference potential of ‘gene’. Later usages of ‘gene’ can either have their references fixed through the description used by Sturtevant in his misclassification (genes are segments within which recombination can’t occur) or through the experiment to which he was responding. My analysis of scientific concepts reveals precisely how, almost from its original usage, the concept of gene was theory laden.

4 BENZER’S REFINEMENT

From 1920 on, the heterogeneity of the reference potential for ‘gene’ became gradually more troublesome. I shall not examine here the numerous debates

1 The case is exactly parallel to Priestley’s extension of the reference potential of ‘dephlogisticated air’, discussed in my [1978].
about the gene concept, debates which can also profitably be reconstructed using the approach to conceptual change offered in 2. Instead I shall simply consider the refinement of the classical concept introduced by Benzer in the 1950s.

If the referent of ‘gene’ is fixed through actual breeding experiments, then there is, as we have seen, a danger that one will assign to the same gene mutation sites between which recombination can occur. When apparent “intragenic” recombinations are then found, one has a choice: the prior division of the chromosome may be retained, at the cost of the recombination criterion, or one may insist that genuine intragenic recombination is impossible, and abandon the previous division. In the late 1940s and early 1950s both options had a following among geneticists. Benzer provided two clear criteria for segmenting chromosomes. One of these uses the principle that intragenic recombination is impossible. The other enunciates a principle of division by the \textit{cis-trans test}, elevated by Benzer into an explicit criterion for capturing that style of segmentation which had hitherto been implicit in the many utterances of ‘gene’ whose reference was fixed through particular breeding experiments. This criterion provided a new means of fixing the referent of ‘gene’ which would preserve the reference of many previous tokens.

The \textit{cis-trans test} can be best explained by imagining a diploid organism in which two mutations occur at different sites within the same functional unit.\footnote{Although Benzer worked with bacteriophage (which are haploid), his test could be applied to such organisms by simultaneously infecting bacteria with different strains of mutant phage.} If the mutations occur on the same chromosome, then the homologous segment will be mutation free, thus allowing for its normal functioning and (let us assume) a normal phenotype.\footnote{The assumption is only introduced to simplify the discussion. As we shall see, what is important is that for pairs of mutations within the same functional unit there’s a phenotypic \textit{difference} between the \textit{cis} form and the \textit{trans} form.} If one mutation occurs in each of the homologous segments, then both segments will function abnormally, yielding an abnormal phenotype. However, if two mutations occur at sites in different functional units then it won’t matter whether they are both on the same chromosome or on opposite chromosomes. To put the points simply, two mutation sites belong to the same functional unit, or \textit{cistron}, if there is a difference in phenotype between organisms in which mutations at those sites occur on the same chromosome (in \textit{cis} position) and organisms in which the mutations occur on opposite chromosomes (in \textit{trans} position). A diagram should make this clear.

# \begin{center}
\begin{tabular}{c|c|c|c}
\hline
\textit{cis} mutation & \textit{normal phenotype} \\
\hline
\end{tabular}
\end{center}

\footnote{1}
Benzer introduced his concept of the cistron to subdivide a genetic region in the bacteriophage T4. Despite the fact that this illustration of the concept was an organism which has primarily been studied by using the tools of molecular biology, Benzer's concept is, as he himself notes, classical. The principle of segmentation into cistrons makes no essential reference to biochemical properties: two mutation sites belong to the same cistron just in case there is a phenotypic difference between organisms with cis mutations at those sites and organisms with trans mutations at the sites. In some cases, the phenotypic characteristics which are of interest (and hence the phenotypic differences) may be biochemical, but this is in no way part of the criterion. Benzer's concept of the gene differs in this respect from the biochemical approach which we shall consider below.

What exactly did Benzer's work accomplish? The development of classical genetics from 1920 to 1950 showed that the reference potential of 'gene' was heterogeneous and cast doubt on some of the assumptions with which this term was laden. By introducing the notion of cistron, Benzer accommodated many of the references of his predecessors, while modifying the reference potential of 'gene'. Prior to Benzer's work, the reference of tokens of 'gene' could be fixed through events in which genes were declared to be entities within which recombination cannot occur or through responses to actual breeding experiments. Benzer provided a description which would characterise the entities singled out in many of the latter events.
and separated the first class of events from the reference potential of 'gene' (or 'cistron'). Geneticists now had a descriptive way of fixing the referent of 'gene' which would accommodate large parts—but not all—of their previous practice. Again, the theory of 2 enables us to state precisely the nature of Benzer's advance, giving a clear sense to the popular idea that Benzer refined the concept of the gene.

5 IMMEDIATE FUNCTION

There was always an alternative way in which to think about the segmentation problem. Once it was recognised that many genes affect some phenotypic characters, it was possible to conceive of those characters as resulting from a complicated sequence of biochemical reactions, isolating genes according to their functional roles at earlier stages of the sequence rather than by the relatively indirect mutation and recombination tests. Thus one might try to recapture a distinct function for each individual gene by focusing on its immediate action, regarding the phenotype as a complex product of such actions. (This idea might even be used to oppose the approach whose history we have traced above: one might maintain that mutational and recombinational tests are inadequate to the unravelling of the functional activity of the genes.) In the late 1930s and early 1940s, Beadle and his colleagues successfully articulated this idea, amassing evidence for the hypothesis that the primary function of genes is to produce enzymes, the 'one gene-one enzyme hypothesis'. This hypothesis could be used to fix the reference of 'gene'. Beadle's work encouraged classical geneticists to initiate tokens of 'gene' by events in which genes were identified as chromosomal segments, each of which directs the production of an enzyme.

The one gene-one enzyme hypothesis needed refining, and developments in molecular biology (the details of which need not occupy us here) have increased our understanding of the primary function of chromosomal segments, replacing the one gene-one enzyme hypothesis with the one gene-one polypeptide hypothesis. I now want to argue that these developments expose a residual problem with Benzer's notion of the cistron and that they show how to resolve many of the controversies which have surrounded the concept of the gene.

In its simplest form, Benzer's division of the chromosome into cistrons assigns two sites for (recessive) mutation to the same cistron if the trans heterozygote is abnormal and the cis heterozygote is wild-type. The sites will be assigned to different cistrons if both cis and trans heterozygotes are wild-type. Now it's important to realise that this test presupposes a conception of the phenotype. Depending on what kinds of properties we count as phenotypic traits, we'll obtain different divisions. But what exactly is a phenotypic trait? Mendel, of course, investigated rather gross features of the organisms he studied; the Drosophila group attended to more subtle effects; Beadle and his co-workers studied phenotypic traits which could be
characterised biochemically. Yet in all these cases, even the last, there is a
gap between the immediate gene action and the phenotype, a gap which
allows for the possibility that abnormal immediate gene functioning might
give rise to a normal phenotype. Molecular biology shows us that the
possibility is genuine.

Our knowledge of the structure of DNA and the nature of the genetic code
leads to a recognition that mutation can arise by means of deletion, insertion,
transposition or simply change of a nucleotide. Consider the last type of
mutation. Whereas deletions, insertions and transpositions are likely to lead
to great variations in polypeptide products, switching one nucleotide will
affect only one codon, thus leading to the formation of a polypeptide with (at
most) only one amino acid difference from the normal product.¹ Although
the abnormal polypeptide may not be able to achieve the full role of the
normal enzyme, it may exhibit partial activity. Thus it is easy to see that
there can be mutants (so-called “leaky” mutants) which approximate (or
even fully realise) the wild-type phenotype at higher levels, even though
they produce abnormal polypeptides. When we apply Benzer’s test to a pair
of mutations within a segment of DNA coding for a polypeptide, both of the
mutations consisting of nucleotide substitutions, we shall assign the
mutation sites to different cistrons if both mutations are leaky. For the trans
heterozygote will produce two abnormal polypeptides, but, ex hypothesi, the
abnormal products will have sufficient enzymatic activity to yield the wild
type. From Benzer on, geneticists have been aware of this possibility, and
there have been a number of responses to it. Benzer’s original studies used
nonleaky mutations in mapping genetic fine structure. For nonleaky
mutants, the phenotype reflects directly the immediate action of the genes so
that segmenting the chromosomes according to Benzer’s test will produce a
division into units each of which codes for a polypeptide. However, if we
once try to arrive at a general classification of genes then, in the case of leaky
mutants, the cis-trans test will divide regions which, as a whole, function to
produce polypeptides. To demand that the trans heterozygote be wild-type
at the level of enzyme production salvages the concept of the cistron at the
cost of severing it from its roots in classical genetics. If to have a wild-type
phenotype is to produce the right enzymes then the concept of phenotype
has become almost redundant. Hence molecular biology shows us that
Benzer’s principle for segmenting chromosomes into functional units will
only tally with the division according to immediate genetic functioning if we
radically distort the classical conception of phenotype.

As I remarked above, the classical concept of the gene presupposed a
concept of a phenotypic trait. What we learn from molecular biology is the
possibility of many different concepts of the gene, generated by different decisions
about the phenotypic level. These concepts may be useful for different areas

¹ Since the genetic code is redundant, some substitutions don’t affect the formation of amino
acids.
of research. We can permit ‘gene’ to have a highly heterogeneous reference potential. Those whose concern is with the continuation of the classical research of Sturtevant, Bridges and Muller, the construction of gene maps for organisms, may reasonably employ the principles of segmentation used by their predecessors. Those who wish to resolve particularly difficult regions, determining genetic fine structure, will find Benzer’s principle of segmentation helpful. Molecular biologists, when pressed to identify the referent of ‘gene’, have adopted, almost universally, the view that genes (more exactly structural genes)\(^1\) are chromosomal segments which code for particular polypeptides. The differences arise from differences about the level at which one wants to define normal genetic functioning, and different decisions are useful for different scientific purposes.

Goldschmidt argued vigorously that there was no correct way to subdivide the chromosome into functional units. In response, defenders of the classical concept of the gene were often inclined to insist that a particular principle of division yields the natural division. The moral of molecular biology is that both parties to the dispute are wrong. Depending on one’s interests, there are various natural ways to segment the chromosomes. The smallest functional unit is the codon, a sequence of three nucleotide pairs. At the other extreme there are lengthy regions of DNA which code for a number of polypeptides whose joint action dictates the form of a readily observed morphological characteristic. In between there are segments which code for polypeptides and segments defined by the application of the cis-trans test at various phenotypic levels. Which of these has the best claim to being dignified with the title of ‘gene’? How can we best obey the maxims Naturalism and Clarity? Well, the units I have discussed are all functional units, and which one you want to isolate as the functional unit depends on the functions you’re interested in. By exposing the structure and modus operandi of the genetic material, molecular biology liberates us from unilluminating controversies. It sanctions an inclusive reference potential for ‘gene’, containing events through which the reference of ‘gene’ may be fixed to meet the needs of different specialties. So long as it is clear that the referents need not always be the same and that none is privileged as the unit of functioning, we can have genes to suit all comers.

6 CONCLUSIONS

Molecular biology has done something important for classical genetics, the genetics of Morgan and Muller, Sturtevant and Bridges, Beadle and Tatum,

\(^1\) Here and elsewhere I ignore the fact that some-chromosomal regions do not code for polypeptides. Anyone familiar with molecular genetics will recognise that the reference potential of ‘gene’ is even more complex than I have taken it to be. In the interests of clarity, the reference of ‘gene’ is sometimes fixed as the set of segments each of which codes for a polypeptide. In the interests of naturalism, the reference of ‘gene’ is fixed as the natural kind including such segments and, in addition, those segments which serve similar biochemical functions (operators and maybe some other segments). Of course, all this is more grist for my mill.
Lewis and Pontecorvo. Philosophers have struggled to explain what the relation between the classical and the molecular theory is. My discussion is intended to show that we need a fine-grained analysis of conceptual change to tell even part of the story. I have endorsed sophisticated versions of three standard historical claims: the concept of gene became theory-laden very early in the development of genetics; at a time when geneticists had come to sense that all was not well with the concept, Benzer refined it; finally, molecular biology exposes and clears up residual difficulties with the concept. By using the general theory advanced in 2, I have tried to explain exactly what these assessments mean.

The example of the development of the gene concept exposes an interesting general point. When one discovers that the reference potential of a predicate permits one to use different tokens to refer to different sets, there are two alternative rational responses. One is to contract the reference potential, eliminating the ambiguity. Benzer’s differentiation of the cistron and the recon is an example of this type of response. The alternative is to acknowledge the ambiguity, recognising the usefulness of fixing the reference of the predicate to different extensions. This latter strategy underlies our response to the enhanced understanding of the genetic material which molecular biology provides. We see that the cis-trans test is incomplete, in that its division of the chromosome can vary according to one’s choice of the level at which ‘wild-type’ if specified. Yet different ways of completing the test can be tolerated, because we know that any problems induced by ambiguity will be resolved if we retreat to a molecular biological description.

It should be no surprise, that, while different modes of reference for ‘gene’ are sanctioned and acknowledged, geneticists now choose to formulate questions in ways which avoid presupposing a particular method of segmenting the chromosomes. Although classical geneticists were often inclined to ask questions of the form ‘How do genes...?’ (e.g. ‘How do genes replicate?’), their modern counterparts are likely to reformulate those questions as questions about the genetic material. Thus Watson and Crick indicated, and Kornberg and his associates elaborated, a solution to the problem of gene replication which presupposes no particular style of dividing the genetic material. Molecular biology is very tolerant: it countenances as many concepts of the gene as the rest of biology may require. Yet, in molecular biological research, talk of genes frequently seems passé, a product merely of the accidents of history. There is no molecular biology of the gene. There is only molecular biology of the genetic material.

*University of Vermont*
REFERENCES


GOOSENS, W. [1978]: 'Reduction by Molecular Genetics', Philosophy of Science, 45, pp. 78–95.


