Genes as Determiners of Cellular Biochemistry

David M. Bonner
Department of Botany and Microbiology, Yale University

THE SCIENCE OF GENETICS was born, and has matured, within the 100 years we are now celebrating. It stems from the acute observations of Mendel nearly a century ago regarding the inheritance of morphological traits in plants. Many years later, T. H. Morgan and his collaborators, working with the ubiquitous fruit fly, gave precise definition to the smallest unit of inheritance, the "gene." Genes are an integral part of the chromosome, arranged in linear order in it, and specific genes are associated with specific hereditary traits. The ingenious work of many investigators has made it possible to map out in considerable detail the relative positions (loci) of a great number of genes on the chromosomes of *Drosophila* and of corn, despite the fact that no gene has ever been seen. We also have extensive knowledge of gene interactions as well as of the gross chemical nature of chromosomes. How a gene determines a specific hereditary character, however, is not as well understood. The progress made in this field in the past several years has been remarkable. There remain, however, great gaps in our understanding of this phenomenon. The mechanism whereby genes can control specific morphological traits, *i.e.* the mechanism of gene action, is not well understood, and we know less still concerning the process that enables a gene to make exact replicas of itself at each cell division. It is the problem of gene action which I plan to discuss at this time.

During the past 20 years considerable evidence has been marshalled by investigators using the fruit fly and corn which suggests that genes probably exert their morphological control through the control of biochemical reactions. In fact, some of the earliest evidence of this relationship stems from the study of hereditary diseases in man, particularly those centering around the metabolism of the two amino acids, phenylalanine and tyrosine. The genealogy of families afflicted with diseases such as alcaptonuria pointed many years ago, toward the genic control of biochemical reactions. Organisms such as man, corn, or fruit flies, however, possess many disadvantages for this type of research. Therefore, in 1941 the geneticist, Beadle, and the biochemist, Tatum, cast about for a better organism with which to make a critical study of gene action and found one ready-made in the bread mold, *Neurospora crassa*. The life cycle of this organism had been worked out by the mycologist, Dodge, and the usefulness of *Neurospora* as a genetic tool had been demonstrated by the geneticist, Lindegren. This fungus is a favorable organism for such investigations. Its nutritional requirements are simple, and it can be grown on a chemically defined medium. It will grow on a medium containing sucrose as a carbon source, nitrate as a nitrogen source, inorganic elements, and the vitamin biotin. From a genetic standpoint *Neurospora* is easily worked with. It exists in two strains of opposite mating type and in its vegetative phase has a single set of chromosomes and genes. Problems of dominance do not therefore arise.

Using this organism, Beadle and Tatum argued that, if genes do indeed control biochemical reactions, it should be possible to induce gene mutations that would affect the synthesis of such vital cellular components as vitamins and amino acids. Such mutations could be detected by the fact that any strain in which such a mutation occurred would then be dependent upon an external supply of this compound. These workers, therefore, subjected asexual spores to agents known to increase the mutation frequency in other organisms, such as X-ray or ultraviolet irradiation. Since the asexual spores are multinucleate, the treated spores were crossed with the opposite mating type of the parental strain to permit segregation of the mutant nuclei. These isolated strains were then established on a complex medium containing yeast extract and hydrolyzed casein as sources of additional growth factors. After the strains were established on this medium, they were transferred to the minimal synthetic medium capable of supporting the growth of the original strain. Most of the isolated strains grew equally well on both media, but an occasional strain appeared which grew well on the complex medium and not at all on the minimal synthetic medium. Using appropriate testing techniques, the factor present in the complex medium and absent in the minimal was then identified, and has proven in most cases to be a known vitamin or amino acid. These occasional strains, therefore, differ from the parental strain in that they require for growth one additional substance furnished in the medium. Included in the list of com-

This paper was presented in the Symposium on Genes and Cytoplasm, held in Washington, D. C., during the Centennial Celebration of the AAAS.

pounds which have been found to be required by various mutant strains of *Neurospora* are all of the B-vitamins, most of the known amino acids, and several of the purines and pyrimidines. That these limited classes of substances have been found in this type of work does not imply lack of genic control of the synthesis of other types of cellular constituents. It reflects, rather, the limitations imposed by existing methods of detection.

As just mentioned, requirements for the various B-vitamins and for numerous amino acids have been found in these mutant strains of *Neurospora*. Is such a biochemical trait hereditary? It is quite simple to perform the proper genetic tests to ascertain the heritability of such a character. For example, numerous strains are known which require either niacin or niacinamide for growth, in addition to the growth requirements of the parental strain. Such a niacin-requiring strain, when crossed with the opposite mating type of the parental strain, gives rise to fruiting bodies which contain sacs (asci) of 8 spores (ascospores). All of these spores, if dissected out and grown separately on a synthetic medium containing adequate niacin, will grow. However, if they are now tested for their ability to grow on the minimal synthetic medium in the absence of niacin, only half of the cultures will grow, *i.e.* half of the spores require niacin, and half of the spores are niacin-independent. The requirement for niacin is inherited, therefore, as a single gene difference between the mutant and parental types. All of the niacin-requiring strains, when crossed with the parental strain, are similarly found to differ from it by alteration of a single gene. Thus, all of these strains represent single gene changes from the wild type. It is of interest to determine whether they all differ from the wild type by alteration of the same gene or whether they represent alterations of several different single genes. This problem is easily settled by crossing the various strains with each other. If the same gene is represented in each case, one would expect that in the spores resulting from the cross only niacin-dependent spores would be found. If they represent different genes, on the other hand, and the genes are not closely linked, one should expect to find an occasional ascus containing some wild type spores, as a result of crossing over. This experiment has been carried out with the niacin-requiring strains, and at least four separate genes have been located which are essential for the synthesis of niacin. Alteration of any one of these four genes gives rise to a niacin-dependent culture.

This type of genetic investigation, though not as yet carried out for all of the known nutritional mutants, has been carried out with a great many of the strains. Each case of a nutritional deficiency which has been examined genetically has been found to be the result of a single gene difference between the mutant and parental types. In general, it has also been observed that several genes are concerned with the synthesis of each of the vitamins and amino acids. It has also been possible to construct a genetic map of the biochemical characters which have been found for *Neurospora*. McClintock has examined *Neurospora* cytologically and has found it to have 7 chromosomes. Using standard genetic techniques for mapping genetic characters, the various biochemical characters have been located on these chromosomes, and we now believe we have at least one character for each of the 7 linkage groups.

The next question to arise is whether the biochemical consequences of these mutations can be analyzed more critically. The investigations of cellular biochemists during the past two decades have been particularly brilliant in exploring the metabolism and synthesis of carbohydrates. Indeed, we now know that a polysaccharide is utilized by a cell in a well-ordered sequence of reactions. Our knowledge of the biosynthesis of vitamins and amino acids has not been nearly so extensive. However, investigations aimed at elucidating the biochemical nature of genetically blocked reactions have proven useful in understanding the biosynthesis of vitamins and amino acids. Biochemical investigations using mutant strains of *Neurospora* have shown that such vital cellular constituents as vitamins and amino acids are also synthesized in characteristic series of sequential reactions, and that the general pattern of synthesis is similar for *Neurospora* and higher plants and animals.

From the existing literature one may draw many examples for illustrative purposes. Although neither of the two series of reactions chosen here are the most completely understood series, both of them will serve to illustrate the different biochemical problems that arise as a result of genic alteration.

As was mentioned earlier, there are known to be at least four genetic types which require niacin for growth. Can these also be broken down into four distinct biochemical groups? Biochemical investigations by Beadle and his collaborators at the California Institute of Technology and by investigators at Yale University have shown that they can. Strains of three of the groups will grow well on niacin and will not grow on the amino acid tryptophane. Strains of the fourth genetic group, however, will grow on either niacin or tryptophane. These are therefore obviously biochemically distinct from the other three genetic types. The latter classes may also be shown to differ biochemically from one another, since strains of one

class will grow equally well on either niacin or hydroxyanthranilic acid, while strains of the other two classes will grow only when supplied with niacin. The two remaining types are differentiated by the fact that strains of one type, when grown on niacin, invariably excrete hydroxyanthranilic acid into the culture medium, whereas the remaining type does not. Thus, these four genetic types appear to correspond to four distinct biochemical types.

Can such biochemical differences be more precisely defined in terms of specific reactions? Several strains of *Neurospora* are known to require tryptophane for growth, and this requirement in these strains cannot be replaced by niacin. These can also be broken into three distinct biochemical classes: class 1, those which can use only tryptophane for growth (*2*); class 2, those which can use either indole or tryptophane for growth; and class 3, those which can use anthranilic acid as well as indole and tryptophane for growth (see Fig. 1). Inasmuch as no mutants have been found which can grow on indole alone, yet cannot utilize tryptophane, indole is established as a precursor of tryptophane. Mutants of class 1 not only differ from the parental type by alteration of a single gene, but they also differ genetically from the other biochemical classes. Biochemically, strains of class 1 have been characterized precisely and are known to differ chemically from the parental type through the loss of a single specific biochemical reaction, viz., the coupling of indole with the hydroxy amino acid serine. Mutants of class 1 have therefore lost the ability to carry out a single specific biochemical reaction by alteration of a single gene. This example is not unique, since many single gene mutations involved in the synthesis of such other compounds as choline, arginine, or methionine are also correlated with the loss of single known reactions. Our present inability to define the reactions blocked in the niacin series, as well as in classes 2 and 3 of the tryptophane series, simply reflects the inadequacy of our knowledge concerning the cellular synthesis of niacin and indole, respectively. All of those strains, however, presumably represent losses of single reactions.

The second series of reactions which I would like to mention deals with mutant strains requiring the two amino acids, isoleucine and valine. A mutant strain has been obtained which differs from the parental type by a single gene change and which requires isoleucine and valine for growth. Detailed investigation has shown that this double requirement is not the result of a common biosynthetic step in the synthesis of these two amino acids. Rather, it has been found that the genetic block prevents the amination of the keto acid analogue of isoleucine. As a result of this block, the keto acid analogue of isoleucine accumulates and in turn competitively inhibits the amination of the keto acid of valine. Thus, the genetic block in one biosynthetic step in this instance results in an apparent double requirement. This example illustrates an important fact regarding these mutant strains. Though the genetic block can be obviated by supplying the necessary end-product, these strains are still not biochemically wild type, since one must still reckon with the fate of the substrates that accumulate abnormally as the result of loss in one biochemical reaction.

Biochemical investigations, in short, show that a gene mutation which gives rise to a growth-factor requirement represents, in reality, a loss of the ability

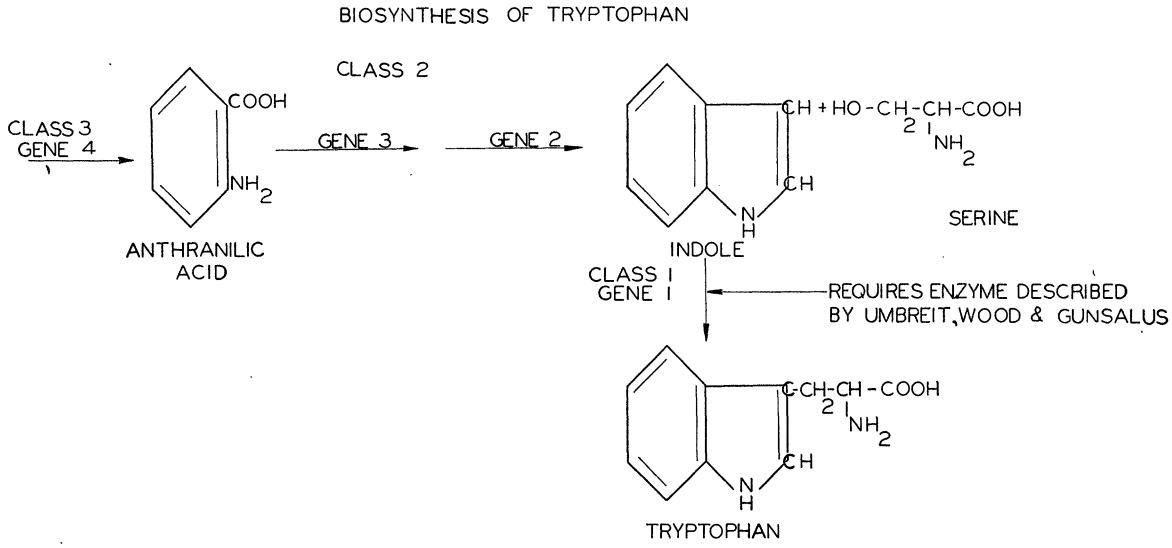

FIG. 1

to carry out a specific biochemical reaction. This view is further supported by the fact that independent mutations of the same gene invariably are associated with loss of the same biochemical reaction. These observations have in turn led to the working concept that single genes control single biochemical reactions. This concept has frequently been referred to as the one-one hypothesis, and we shall return to it later.

Having now seen the evidence for assuming that single genes control single biochemical reactions, can we probe more deeply into the nature of this control? It is here that the evidence becomes very much weaker. It has often been postulated that genic control is exercised through control of enzyme specificity. Since nearly all biochemical reactions require enzymatic catalysis, the assumption that the genetic control of biochemical reactions is through control of enzyme specificity is attractive and plausible. Direct evidence of any such control is still scant, though several lines of evidence do support this point of view. If gene mutation does give rise to enzyme alteration, one might expect substrate accumulation. This has been observed many times, one example being the accumulation of hydroxyanthranilic acid by a *Neurospora* mutant. However, this is at best only indirect evidence. More convincing would be the direct isolation of an altered enzyme. Investigations at Yale University have recently been aimed at this problem. *Neurospora* is capable, under proper environmental conditions, of adapting to utilize lactose as a source of carbon. It has been possible to isolate strains, differing from the wild type by alteration of a single gene, which under similar environmental conditions cannot utilize lactose as a carbon source but will grow on sucrose. An enzyme has been prepared in cell-free extract from the parental type which will split lactose to glucose and galactose. In preliminary experiments this enzyme has not been found in the mutant. In this instance, then, the loss of a specific biochemical function appears to be associated with mutation of a single gene and also with lack of production of a specific enzyme. Recently, Mitchell and Lein (*2*) have made similar observations regarding the enzyme, described by Umbreit, Wood, and Gunsalus, for the condensation of indole and serine to the amino acid tryptophane. Mitchell and Lein have found this enzyme lacking in those mutants, discussed earlier, which require tryptophane for growth and which are unable to replace this tryptophane requirement by indole. Here, then, is a second instance in which a gene alteration leading to the loss of a specific biochemical function is associated with the absence of a specific enzyme. These observations, while aiding the argument that genes do control enzyme production, do not contribute critical proof of genetic control of enzyme specificity. The correlation of altered gene to altered enzyme, rather than simple loss of enzyme, is yet to be established.

The biochemical and genetic observations which have just been reviewed lead to the conclusion that single genes control single biochemical reactions, and that this control is exercised through control of enzyme production. These conclusions have given rise to the thesis that a one-one relationship exists between gene and biochemical reaction, and that this is the reflection of a more fundamental relationship existing between genes and enzymes, *i.e.* that single genes determine the specificity of single enzymes. There is considerable logic in this point of view, since the evidence at hand suggests that genes are nucleoproteins. The possibility that the gene serves, therefore, as a template in determining the final configuration of the enzyme is an attractive theory and has been considered by several investigators. The thesis of a one-one relationship of gene to specific enzyme has, however, been justifiably criticized, and, as Delbruck has pointed out, the experimental methods which have been used may have largely precluded the detection of incompatibilities. Investigations at Yale University aimed at the general problem of gene-enzyme relationship as discussed earlier have shown that genic alteration leading to the inability to utilize lactose as a carbon source is correlated with loss of the enzyme, "lactase," the enzyme which splits lactose to glucose and galactose. More interesting, however, is the fact that alteration of any one of two and probably three independent genes results in the loss of lactase production. The production of the enzyme, lactase, appears, therefore, to be controlled by more than a single gene. Lederberg (*1*), working with the bacterium *E. coli,* has also recently observed that alterations of any one of 7 independent genes affects lactase production in this organism. These preliminary data suggest that enzymes, as other cellular constituents, are synthesized by a characteristic sequential series of biochemical reactions. The various characteristic biochemical reactions in the synthesis of a given enzyme undoubtedly are enzymatically catalyzed. Such considerations, however, emphasize the difficulty attendant in a thesis such as a one-one relationship of gene to specific enzyme, since it is difficult in our present state of knowledge to decide whether enzyme specificity is a property conferred upon an enzyme in the final stage of synthesis or whether specificity represents the logical culmination of its entire characteristic synthesis. How might one then sum up current thoughts concerning the relationship of gene to biochemical reaction? Single genes are known to control single biochemical reactions, and this control is probably exerted through control of enzyme production. The nature of the relationship existing between gene and enzyme remains a major biological enigma.

One might next logically question whether the theory that genes control enzyme production fits in suitably with known genetic relationships. It is obvious that it cannot account fully for all of the known complicated biological data. Our knowledge regarding the more intimate structure and properties of proteins is scant indeed, and with genes one is likely dealing with the most complicated type of protein, a nucleoprotein, a self-duplicating particle. Many genetic phenomena are, however, certainly accounted for in part by such a theory as, for instance, multiple allelism. If genes do control enzyme production, it is possible that various alterations of a given gene might lead to enzymes of various types or to an enzyme entirely lacking specificity for its customary substrate. On the other hand, a less drastic gene alteration might give rise to an enzyme possessing specificity under certain environmental conditions and lacking it under others. There are, for example, mutant strains of *Neurospora* which are different from the wild type only under certain conditions of temperature or pH, while allelic strains also occur which are mutant in phenotype under all conditions of temperature and pH. Many other sorts af alleles could probably be found if suitable methods were available for detecting them, e.g. forms showing altered efficiencies, forms converting the substrate to different end-products.

In this discussion the relationship of gene ⟶ product ⟶ morphology has been omitted. Morphological changes have been repeatedly observed after the irradiation of *Neurospora.* Inositol concentration can determine whether growth will be diffuse or colonial, and recent work has shown that numerous compounds can produce effects which simulate that of a single gene known to control growth habit. These phenomena cannot, however, be precisely defined biochemically. That decided morphological changes might be occasioned by subtle enzyme changes is fascinating speculation, genetic evidence for which Stern has now presented.

Nearly every biochemical reaction that occurs in a living cell requires enzymatic catalysis. Since genes are necessary for the production of each of these enzymes, the question is often raised: Does not a cell require an astronomical number of genes, inasmuch as the number of cellular reactions is undoubtedly large? Although an answer to this question must be largely speculative, the order of magnitude of the number of genes estimated by geneticists to be present in various organisms is 10,000. Are 10,000 enzymes sufficient to run a cell? I suspect so. It is of importance to remember that enzymes have relative specificity, and that tremendous variation in structure of large molecules such as proteins can theoretically be achieved by relatively few enzymes. Furthermore, if one adds up all of the reactions necessary to synthesize all of the vitamins, amino acids, purines, pyrimidines, and all of the reactions necessary for the metabolism and synthesis of polysaccharides, the total is still only of the order of 1,000, leaving scope for at least 10 times as many other sorts of reactions. Thus, at present no serious disparity between the number of reactions vital for a living cell and the number of genes that such a cell probably contains is apparent.

This discussion has dealt primarily with the one organism, *Neurospora.* This has been deliberate, since at the present time *Neurospora* presents in many respects the best material for correlating genic and biochemical changes. Many important speculations and investigations resulting from the use of biochemical mutant strains of microorganisms have been omitted, and instead the emphasis has been placed on gene action. Genic control of biochemical reactions undoubtedly holds equally well for all other organisms; it has, in fact, been demonstrated in a great many. There is little reason to believe that the mechanism of gene action in microorganisms such as bacteria and fungi is substantially different from that of macroorganisms, such as insects, higher plants, and mammals. *Neurospora* has served well for relating genes to biochemical reactions, but it may well not be the best possible material with which to tackle the problems that the future holds for us in this field. The relationship of genes to enzymes will undoubtedly be different from that which we envisage at the present time, but will probably prove even more exciting.

References

1. LEDERBERG, J. *Rec. genet. Soc. Amer.,* 1948, **17**, 46.
2. MITCHELL, H. K., and LEIN, J. *J. biol. Chem.,* **1948, 175**, 481.