

Lecture #46 - Virus Genetics: Bacteriophage  
by J. Lederberg

So far our discussion has emphasized behavior of organisms of a considerable degree of complexity. These organisms, starting with bacteria, including *Drosophila* and mammals, have a sufficient content of information in their genetic material so as to specify a large number of different enzymes. These enzymes, together with other proteins which make up various structural features, suffice to give these organisms a free-living character. A bacterium, for example, is capable of subsisting in a relatively simple environment which it converts to its own use through a wide range of different synthetic pathways. For each step of these pathways we must consider the probable presence of a distinctive protein whose amino acid sequence has been specified by its DNA content.

Probably the most significant measure of biological complexity of different organisms is furnished by a count of their nucleotide information expressed in the pairs down the length of the postulated Watson-Crick helices in the DNA. And we see that a bacterium, for all of its small size, already has 6 million nucleotide units per nucleus as compared with the 80 million in a *Drosophila*, and the 5 billion in a mouse nucleus. If we assume, as now appears to be an approximately correct estimate, that it will require something of the order of 2000 nucleotide units to specify a single protein of average size, we see that there is sufficient DNA in a bacterium to specify something like 3000 different enzymes. Now by no means no number of enzymes remotely approaching this number has actually been isolated from any single bacterium, but if one considers the wide range of metabolic activity which free-living organisms are capable of sustaining, this is by no means an unreasonable estimate. In fact, one might be sur-

prised that a bacterium could get along with so low a number.

When we come to the simpler organisms that will be the subject of the current presentation, we see that their genetic simplicity already poses a definite limitation on the possible complexity of biological processes that they are capable of sustaining. The phages that we have been discussing so far are particles whose nucleotide content is of the order of one hundred thousand pairs, and this would allow, according to the same estimates, for the production of something like 50 different proteins. Considering the complexity of the enzyme systems that would be necessary for the production of the amino acid constituents of these proteins and for their assembly and for the production of the elements of the DNA and its assembly in the bacterium, we would have to conclude that any particle of these dimensions must rely very heavily on the available metabolites, and not only metabolites, but the metabolic systems of the host in which this particle resides.

Given this degree of simplicity, of a bacteriophage like T2 or lambda, it is therefore not at all surprising that these organisms are incapable of a free-living existence, that they are unable to multiply and produce more of their kind except within the confines of a metabolically active host cell. This host cell provides most of the accessory metabolism which is necessary for the synthetic processes of the virus; the virus adds those elements which are unique to its own structure, enabling it, on the one hand, to compete with the metabolic systems of the cell that it infects, and, on the other hand, to produce protective structures which will enclose the virus during the interval when it must live outside the host cell.

When we come to the smaller viruses, for example, the animal viruses in the encephalitic diseases, and an exceptional bacteriophage called phi X 174, we find that

we can have living particles of even smaller dimensions, of the order of 5000 nucleotide units, which is hardly more than the information needed to, which is encompassed within, one or two, something of that small number, of distinctive genes. It is of some interest that these smaller viruses so far established are all one-stranded in their nucleic acid composition. Most small viruses, and this is true of all of the small viruses of plants, and of animals as well, are composed of RNA which is typically one-stranded, or whose structure at least is not so decisively determined in the two-strand configuration as is that of DNA. The virus just mentioned, phi X 174, has created some excitement in the world of biochemical genetics since it has recently been found that it is a unique example of a DNA configuration which is one-stranded. In fact, some authors have suggested that this virus nucleic acid really is much more reminiscent of RNA than of DNA because of its secondary structure, although it has all of the basic elements, of a deoxyribose sugar, and of the characteristic bases, that we see in other DNA molecules.

We must consider that a virus of this sort does a very remarkable job since it is capable by virtue of the production of perhaps two distinctive proteins, or something of that order, of sufficiently diverting the normal metabolism of the host cell which it infects so as to result in the production of large quantities of this particular nucleic acid, and of the protein associated with this nucleic acid, rather than the normal metabolic products of its host.

Now if we look a little bit more closely at the structure of a bacteriophage particle, bacteriophage being a rather special class of viruses with a structure unique to these, these are the viruses attacking bacteria, we can get a schematic diagram, something like this. We are dealing here with a particle whose characteristic dimension is about one tenth of a micron, about a tenth the diameter of the bacterium that it in-

fects. It has a protein coat which has a somewhat crystalline form; within that protein coat is a tightly packed unit of nucleic acid. In fact, one can calculate that, if properly extended in a double-strand helix, this nucleic acid will be some 34 $\mu$  long, and yet it is packed in an efficient and orderly way in a box of these particular dimensions.

Accessory to this protective box around the DNA is a tail, which has a spiral protein around it which may serve some contractile function in the penetration of a host bacterium by a bacteriophage, and there are tail tip fibers which are believed to play an important role in the specific adsorption of the bacteriophage particle to the bacterium.

This structure which is characteristic of the virus in its free state, (in) outside the bacterial cell, is however not found in the recently infected bacterium. And in fact for many years there was considerable mystery concerning the basis of the so-called eclipse period immediately after the infection of a host cell when a bacterium that was destined again to yield intact bacteriophage did not demonstrably contain any element of it. This mystery was eventually resolved by tracer labeling experiments by Hershey and Chase, who found that they could differentially label the nucleic acid component of the bacteriophage particle by incorporating radioactive phosphorus 32 into the DNA as against the protein component which contains sulphur amino acids and could be labeled with radioactive S-35. They found that after bacteriophage had adsorbed onto sensitive bacteria, and electron micrographs support the contention that this adsorption is first at the tail, that there was subsequently a separation of the P-32 labeled DNA from the protein. The P-32 somehow enters the bacterium, leaving a shell of protein outside. In fact, Hershey and Chase were able to demonstrate that after the bacteria had been initially infected they could strip off the empty shells containing all or virtually

all of the protein of the bacteriophage by exposing the bacteria to high shearing forces, as in a Waring blender. This, however, did not in any way alter the fate of the inoculated bacterium which had received the genetic element, the DNA from the bacteriophage. However, this is not, this DNA is not, a bacteriophage particle. This is the genetic element of the bacteriophage, but before a particle is produced which has the capacity for free existence outside the cell, and which is capable of reinfecting another cell, a new cycle of protein synthesis and organization and assembly of the intact phage particle must first take place. This does not begin for several minutes after the initial infection of the bacterium. During this time there is considerable replication of the DNA component, which is now called the vegetative phage to distinguish this from the intact phage particle, which is the infective phage particle. This vegetative phage is considered to form a pool of DNA units of the replicating DNA of the phage from which from time to time components are sampled, in the sense that a fraction of the vegetative pool will undergo a condensation of the DNA and then around that DNA a new skin, let's say a new head and a new tail, will be produced, again to form an infective phage particle.

As a consequence of other materials produced by the bacterium which is infected by the bacteriophage, namely endolysins, which are enzymes that can lyse a cell wall, after some 20 to 30 to 40 minutes the bacterium will be lysed and the infective phage particles contained within the bacterium will be liberated in the medium. These infective particles are stable in free suspension, and can persist until the next occasion when they encounter a suitably competent bacterial host, again to initiate a cycle of infection.

This is the typical cycle for a lytic bacteriophage, or to speak more precisely, this is the lytic cycle of a bacteriophage, and it represents one of two alternatives which is available to a bacteriophage like lambda. As we already know, lambda is capable of

penetrating a bacterium, and it does so through a separation of its DNA and protein in much the same fashion, but here the vegetative lambda can pursue either the course of lytic multiplication resulting, that is, in the ultimate lysis of the bacterial cell, or in establishing a new relationship with the chromosome of the bacterial host, in this case giving rise not to a dead, lysed cell, but to a lysogenic bacterium, which betrays its peculiarity by conferring the property on its descendants of eventually producing some of the same bacteriophage.

Now, when this bacteriophage is liberated into the medium it not only destroys the cell from which it has been liberated, but it is also capable of now, these new particles are capable of, infecting any bacteria of a susceptible type that may be their environment. Consequently, we can readily detect the presence of a bacterial virus by using a sensitive strain as an indicator host. This can be done in liquid medium, in a fairly obvious way, by waiting for the lysis of the entire bacterial culture. However, for many purposes of enumeration it is more convenient to conduct these assays on a solid medium. These assays represent one instance of the great utility of looking at nothing to get a considerable amount of information.

What we set up is a Petri dish with an ample supply of nutrient agar, and on top of this dish we add a layer which is heavily seeded with sensitive host bacteria. Each of these bacterial cells will form a small clone, as you already know, but the number is so large that they tend to merge into one another, and we have what we may very appropriately call a lawn of bacterial growth. If, at the same time as the bacteria are added to this layer of agar, we have added a few phage particles, they will find themselves in the vicinity of some susceptible bacterium, the phage particle after one cycle of lysis will result in the liberation of some hundred, or several hundred, daughter particles,

which will then diffuse to the neighboring bacteria in the vicinity of the original burst. This will then result in a progressively increasing zone of lysis, or of clearing of the bacterial growth, at the point where a bacteriophage particle has alighted. I have an example of that kind of a plate here, and this plate, which will now be projected on the screen, represents a plating of the bacteriophage lambda, of which I've already spoken. I would just ask you to disregard that white disc spot in the middle, which is a bacterial contaminant that has overgrown in the middle of the plate. In these clearings, which you can see in the plate, are each of them so to speak, a colony of lambda which has grown at the expense of the bacteria in its immediate environment. At the same time as the lambda has grown, the other bacteria that have not been exposed to lambda, have also grown, and they give rise to the opacity against which the colony of lambda, the plaque, appears as a clearing. We therefore have a ready way of counting the number of bacteriophage particles that were present in a given sample added to the bacteria on this lawn.

In addition there are characteristic features of the plaque, which we cannot always trace to any very simple attribute of the bacteriophage-host system, but which do give us the basis of a morphological differentiation of different kinds of plaques, and hence of the genotypes of the phages which produce those particular plaques.

Now we may very appropriately inquire what there is about the structure of the DNA of the vegetative bacteriophage which has infected a bacterium, and which we know to be multiplying very rapidly, that distinguishes it from the existent DNA already present in the host cell. Now we must assume that since we do have bacteriophages which are quite virulent, which are capable of immediately stopping the synthesis of other proteins, and indeed of the synthesis of bacterial DNA, that there is some remarkable element in the chemical structure of the virus DNA to account for this. In at least

one case we have a clue as to the basis of this differentiation, because the phages T2, T4, and T6 do indeed have a characteristic base, hydroxymethyl cytosine, which appears in the DNA of these phages in place of the cytosine, which is characteristic of most other DNA. And it can hardly be a coincidence that hydroxymethyl cytosine, as far as we know, is found only in these particular bacteriophages. And it appears likely that the hydroxymethyl cytosine DNA has features of resistance to the degradative enzymes which will be present in the bacterial cytoplasm, which will account, at least in part, for the unique reproductive capacity, which is to say the virulence, of this DNA as an intracellular virus.

However, we cannot appeal to this type of peculiarity in chemical composition to account for the virulence of other phages like T1, or lambda, where we can find no detectable differences in base composition. This dilemma is, in fact, especially pointed for temperate phages like lambda, which can function as part of the bacterial genome in the lysogenic cycle, and then can be liberated, for example by ultraviolet induction, to function as lytic bacteriophages. We are left here with a mystery which will require considerable further investigation of the biochemistry of virus infection to elucidate.

Having established the basic features of the life cycle of a virus, we may now inquire of some aspects of its genetic behavior. In order to demonstrate genetic interaction between virus particles, it is, of course, necessary to doubly infect a single host cell with two such particles. This can be done by simply mixing sufficient bacteriophage with a suspension of bacteria to insure that each bacteriophage will have simultaneously adsorbed two or more bacteriophage particles -- I said bacteriophage, I meant each bacterium adsorbed two or more bacteriophage particles. If these are of

different type we then have the opportunity to study the interaction between vegetative DNA derived from bacteriophage particles inhabiting the same bacterium.

What has been found, as one might anticipate from the theme of all of these lectures, is that, quite regularly, these mixed infections result in an exchange of genetic material. This can be demonstrated, for example, by the use of such markers as the host range quality, and the plaque type quality, which are designated by the symbols  $h$   $r$ . If one bacteriophage is a double mutant, another the wild-type (and the plaque type with respect to) -- wild-type with respect to plaque and host range -- we then can find in the issue from such mixedly infected bacteria not only the two parental types, but also the two recombinant types,  $h^+ r$  and  $h r^+$ .

The simplicity of the DNA structure of bacteriophage compared to other organisms has encouraged the hope that it might be possible to approach the ultimate complexity of the DNA in genetical experiments. We would, after all, in principle, in order to completely map phage T2, have only 100 thousand different mutants at different sites at our disposal. Now, we have not approached anything quite like that level of precision in the analysis of the entire genome, but Benzer has set out to do this in phage T4, with respect to a definite restricted part of the total genome of that phage.

In phage T4 there is evidence for a single linkage group of the extent of 100 thousand nucleotide units, within which there is a segment whose length approximates 1% of the total map length of the bacteriophage, and this segment is labeled  $r$ . A large variety of  $r$  type mutants have been found in T4 by Benzer and classified by attempts to cross them with one another. Fortunately an easy selective system was available, whereby  $r$  mutants can be distinguished from the wild-type, since the wild-type being capable of growing both on cells of strain B and of strain K12 of *E. coli*, and the mutant being capable of B, but not of growing on strain K12, it was an easy task to isolate

rare particles issuing from mixed infections of r mutants which had reacquired the capacity to grow on strain K12.

In this way it becomes a simple task for Benzer to study the recombination occurring between bacteriophage particles in mixed infections at rates of as low as  $10^{-6}$  or  $10^{-8}$ . By studying a large collection of r mutants Benzer has found it possible to classify about 1500 of them into some one hundred different sites within this r segment of the bacteriophage. The implication of this is that there have been many recurrences of mutations at certain sites, and indeed it does appear that there are some so-called "hot spots" which are much more liable to undergo spontaneous mutation than other sites within this particular locus.

Now, considering that 1% of the bacterial map should give us approximately a thousand nucleotide pairs, a thousand units susceptible of mutation in the locus, it is apparent that Benzer has not yet quite exhausted all the mutational possibilities in this region. However, he has achieved a reasonable step in that direction, and he does indeed find that the distance of, that the frequency of least recombination which has been found between, any single pair of mutants is of the order of .01%, which is in close agreement with that predicted from the dividing the total map length of about 10 morgans by the length of the segment that we are dealing with here. He therefore concludes, with some degree of plausibility, that the minimum distance of recombination, observed genetically, corresponds to the least unit for which we can attribute any chemical basis, namely the interval between two successive nucleotides in a DNA string.

Benzer can also study the phenotypic interaction between different mutant phages growing in the same bacterium. A single bacterium may be mixedly infected with a mixture of two distinct r mutants. This affords an opportunity to determine what possi-

bility, if any, there is of phenotypic cooperation between the vegetative DNA of these two genotypes. Benzer found that many combinations of r mutants were capable of cooperating so as to result in the growth of each of them and the lysis of the cell that contained them, although by definition the individual particles of  $r_1$  or  $r_2$  would be incapable of effecting lysis of the K12 cell. The yields from such mixedly infected cells in these cases would consist of particles, for the most part, either of the parental  $r_1$  or of the parental  $r_2$  type, and the phenotypic cooperation should be carefully distinguished from recombination which gives rise to genetically wild-type particles.

A complete classification of Benzer's mutants resulted in the establishment of two groups within the r region. These groups may be considered cistrons, since the test for cooperation of the two genotypes within a single bacterial cell is equivalent to the cis-trans test of cooperation in a heterozygote. It was found that all of the mutants could be classified either into cistron A or into cistron B, so that the mixed infection of a single bacterium with one member of A and one member of B resulted in effective cooperation. From this Benzer concludes that there are two separate protein elements synthesized under the governance of this region, both of which are required for the effective growth of a bacteriophage in this strain.

Further evidence as to the fine structure of the genetic material is furnished by analyzing a collection of mutants with a somewhat different property from the point mutants now given. These behave as if they are deletions covering various parts of the total  $mpe$ , indicated by these segments over the point mutants on the diagram. Benzer found that one could uniquely assign the position of the point mutants in one and only one order so as to give a compatible superimposition of the various deletion types. Consequently we can conclude that even in the utmost details of its fine structure the genetic material of phage T4 is organized in a linear fashion.