THE EVOLUTION OF GENE CLUSTERS AND GENETIC CIRCULARITY IN MICROORGANISMS[1]

FRANKLIN W. STAHL[2] AND NOREEN E. MURRAY[3]

Medical Research Council Unit for Molecular Biology, Hills Road, Cambridge, England

Received November 2, 1965

THE striking degree of clustering of genes (i.e., cistrons) governing similar characters in T-even phages (and in the Enterobacteriaciae) has been repeatedly remarked upon (e.g., EDGAR and EPSTEIN 1965a, b; DEMEREC 1965). The operon concept (JACOB and MONOD 1961) provides a rationale for those gene clusters in which coordinated gene expression has been established. However, it sometimes appears that this valuable concept dominates the thinking of microbial geneticists to the exclusion of a highly reasonable notion about gene clusters which has historical precedence. FISHER (1930) stated conditions which would confer a selective advantage to an increase in the degree of linkage of genes with related functions. His basic idea has been elaborated by others (e.g., LEWONTIN 1965). For *our* purposes, FISHER's argument can be paraphrased as follows.

For any population the presence of genetically variant individuals potentially adapted to various possible environmental changes promotes the survival of the population. Such variants can arise directly by mutation. More importantly, perhaps, they can arise by recombination among polymorphic genes. Such recombination will result from the encounter (mixed infection in the case of virulent phages) of two individuals who are themselves well adapted to the environment of the day (i.e., quasi-wild-type individuals—a T2 and T4 particle, for instance.) Mixed infections between different T-even phages are certainly far less frequent than genetically unmixed infections. If these relatively rare interactions are to provide potentially useful progeny, the variation must be released by recombination in a sufficiently conservative fashion to ensure that *some* more or less viable particles are produced. These considerations impose a positive selective value upon linkage arrangements which minimize recombination between genes whose products most intensely interact in determining viability in the environment of the day. Following a suggestion by G. STENT, we shall call such a linkage arrangement an "orthotopic linkage system." The clustering of genes controlling the same character provides *prima facie* evidence of an orthotopic linkage system.

BODMER and PARSONS (1962) have previously suggested that some cases of close linkage of genes with related functions in microorganisms have evolved for such

[1] Supported in part by a research grant from the National Science Foundation (GB-294).

[2] On leave from the Institute of Molecular Biology and Department of Biology, University of Oregon, Eugene.

[3] On loan from the Botany School, Cambridge University. Supported by a NATO fellowship from the Department of Scientific and Industrial Research (Great Britain).

populational as opposed to purely functional reasons. THOMAS (1964) has made the specific suggestion that populational considerations may account for the tight linkage between genes governing immunity in the lambdoid phages. The experiments described herein using T-even phages were performed in an effort to provide additional illustrations of orthotopic linkage systems in microorganisms as defined by three criteria—(a) the products of some pairs of genes interact strongly in determining viability; (b) such pairs are polymorphic among the T-even group of phages; (c) the members of such pairs of genes are closely linked. These three criteria are simultaneously demonstrated by any case in which complementation between a pair of closely linked genes, while good when both complementing phages are either T2 or T4, is poor when one of them is T2 and the other is T4. Such an observation (with proper controls) establishes three conditions necessary for an orthotopic linkage system. It does not, of course, demonstrate that orthotopic considerations do in fact account for the observed clustering.

Outline of Experimental Approach

Amber mutants of phage T4 are (by definition) unable to grow on *E. coli* B (EPSTEIN *et al.* 1963), though they can be propagated on those *coli* strains which suppress the amber phenotype. Amber mutations can occur in many parts of the T4 genome (EPSTEIN *et al.* 1963) and in *coli* B result in the production of truncated polypeptide in place of the normal product of that genic region (SARABHAI *et al.* 1964). The amber mutants of T4 have been assigned to cistrons by a qualitative spot test (EDGAR *et al.* 1964). Since complementation between amber mutants is typically "all-or-nothing", this simple procedure reliably determines whether or not two ambers are in the same cistron.

We isolated 84 amber mutants from T2L. We then assigned them to complementation groups (presumably cistrons) by spot-testing them against each other. One representative from each T2 complementation group was then spot-tested against a T4 amber from each of many of the presently identified cistrons. Many T2 mutants gave spot-tests which were clearly positive in every case except one which was clearly negative. "Clean" responses of that sort permitted the establishment of homology between a T2 complementation group and a particular T4 cistron. Some mutants, described more fully below, gave complex results in the spot-tests. Quantitative complementation tests plus relevant controls (see *Details of Experimental Approach*) were then performed between T2 and T4 ambers from different cistrons. The operational definition of interaction of products of polymorphic genes is the observation that an intertype complementation test gives significantly fewer phage particles than do the homologous intratype complementation tests.

Details of Experimental Approach

Materials. Phage strains: Amber mutants of T4D were from the collection of S. BRENNER; most of them were originally isolated by R. H. EPSTEIN. Amber mutants of T2L were isolated from a wild-type stock grown in 5-bromodeoxyuridine by S. BRENNER.

E. coli strains: Strain BB (obtained from J. W. DRAKE) was used as host for the quantitative complementation tests. Strain B was used for spot-tests. Stocks of T4 amber mutants were made on CR63. Stocks of T2 ambers were made on 011′, a permissive derivative of B obtained from S. BRENNER. All platings in the quantitative complementation tests were made on 011′.

Methods. In general, the methods were those described by ADAMS (1959).

Spot-tests: Plates were prepared in which the top layer contained about 2×10^8 B cells and 5×10^6 particles of a particular amber mutant. Spots of each of a number of other amber mutants were then placed in a nonoverlapping manner on the surface of the solidified top layer. The titer of phage in the suspensions spotted was 5×10^7/ml. Generally, the spots obtained after overnight incubation were either turbid or contained about 100 plaques indicating respectively no complementation or complementation and recombination.

Quantitative complementation test: E. coli strain BB was grown in broth with aeration to 2×10^8/ml. Cyanide was added to a final concentration of 0.004 M. Equal volumes of cell suspension and phage suspension containing 10^9 particles/ml of each of two genotypes were mixed. Since over 90% of the particles were adsorbed, the total multiplicity of infection was about 10. After 20 minutes at room temperature without aeration, the infected cells were diluted 10^4-fold into broth at 37°. Unadsorbed phage were determined by assaying a portion of the infected cell suspension chloroformed 10 minutes after dilution. After 70 minutes at 37° the diluted cell suspension was chloroformed and the phage yield titrated. Along with each set of complementation tests the yields from the following mixed infections were measured: (1) T2 wild type with each of the relevant T4 ambers and (2) T4 wild type with each of the relevant T2 ambers.

RESULTS

Spot-tests between T4 ambers and T2 ambers: An amber representing each of our T2 complementation groups was spotted against each of the T4 ambers indicated in Figure 1. Many of the spot-tests were clearly positive or negative. A number of them were ambiguous for easily imaginable but unknown reasons. In Table 1 are presented the results of all the possible pairwise spot-tests. Three cases looked particularly promising. T2 complementation groups B and X both failed to give positive spot-tests against either cistrons 37 or 38 of T4; T2 group F failed to give positive spot-tests against either 43 or 44. The failure to give a positive spot in this test can result from either of two causes—a failure to complement or a failure to recombine to yield amber$^+$ particles. Quantitative complementation tests, therefore, were performed with these mutants. For T2 group F, the spot-test was shown to be misleading; group F, while failing to complement an amber in cistron 44, complemented well with an amber in cistron 43. The results of the quantitative complementation tests with cistrons 37 and 38 versus complementation groups B and X, which map in homologous positions (R. RUSSELL, personal communication), however, provided data which illustrate the sought for phenomenon.

Quantitative complementation tests with genes 37 and 38: The intratype complementation tests (Table 2) resulted in burst sizes comparable to the burst sizes encountered by us for T4 and T2 wild-type phage. For T4 the wild-type burst size was 251 (average of five experiments whose extreme values were 220 and 275), and for T2 the burst size was 66 (average of two experiments whose values were 74 and 57). Mixed infections with T2 amber B mutants and T2 wild type produce less phage than do infections by T2 wild type alone. This partial dominance of T2 amber mutants of group B presumably accounts for the difference

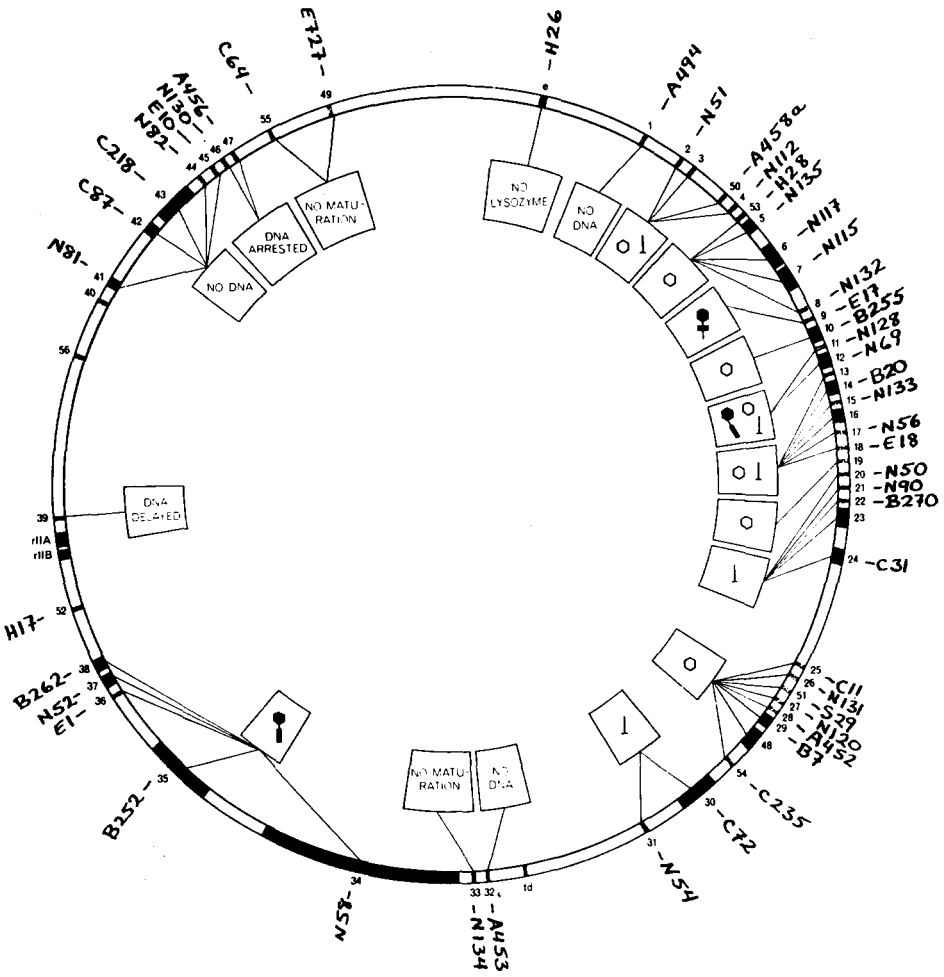

FIGURE 1.—The map of T4 according to EDGAR and EPSTEIN (1965b). The T4 mutants employed in this study are indicated on the outer limits of the figure. The circle within that shows the numerical designation of the genes. The innermost symbols relate to the phenotype of particles mutant in the indicated genes (see EDGAR and EPSTEIN 1965).

in burst size between T2 wild-type infected cells and cells mixedly infected by B and X.

The intertype complementation tests gave essentially no progeny in each of the four possible combinations. The results of the wild type by amber mixed infections rule out the possibility that either T2 or T4 is rendered nonfunctional for the genes in question in the presence of the other. The fact that T4 ambers in genes 37 and 38 are complemented in the spot-tests by all T2 ambers except B and X provides qualitative assurance of the same sort.

Other quantitative complementation tests: Other cases of interaction of products of linked genes were sought by quantitative complementation tests involving a

TABLE 1

Results of complementation spot-tests of T2 ambers vs. *T4 ambers*

T2 amber	T2 complementation group	T4 gene numbers for which test was other than clearly positive
2	A	? on 49
4	B	— on 37, 38
5	C	— on 30, ? on 34, 41, 42, 43, 45, 46, 49
6	D	— on 35, ? on 42, 43, 45
7	E	? on 32, 43
8	F	— on 43, 44
9	G	? on 7, 32, 34, 42, 43
10	H	none
13	I	— on 12, ? on 32, 43
15	J	none
18	K	none
19	L	none
23	M	— on 36
29	N	? on 42, 43
32	O	— on 46
40	P	? on 34
43	Q	? on 32
45	R	? on 32
48	S	none
51	T	none
56	U	— on 55
62	V	— on 14
63	W	none
64	X	— on 37, 38
66	Y	— on 54
44	AA	— on 33, ? on 32
57	BB	— on 27
78	DD	? on 43
82	EE	— on 51
83	FF	— on 42
87	GG	? on 43, 49
98	II	— on 45, ? on 42, 43
99	JJ	none
109	KK	none
114	LL	— on 4
115	MM	— on 43

Each of the indicated T2 ambers was spotted against all of the T4 ambers indicated in Figure 1. Most such spot-tests showed clear evidence of complementation; in the Table are entered all those cases which did not do so.

number of pairs of neighboring genes. Occasional cases of poor intertype complementation were observed. In all these cases, however, the intratype complementation was equally poor. Comparisons of the yield obtained in intratype *cis-trans* comparisons revealed the existence of *cis-trans* position effects for these "cistrons". A report on these experiments and others related thereto will be prepared soon.

TABLE 2

Quantitative complementation tests involving T4 genes 37 and 38 and T2 complementation groups B and X

Test	Burst size*			
	Expt. 1	Expt. 2	Expt. 3	Expt. 4
T4 37 × 38	264	248	..	..
T2 B × X	22	40	..	..
T2 B × T4 38	<1	3	..	..
T2 X × T4 38	<1	<1	<1	<1
T2 B × T4 37	<1	<1	..	..
T2 X × T4 37	<1	<1	<1	<1
T2 wild × T4 37	14	10	13	13
T2 wild × T4 38	20	18	24	18
T4 wild × T2 B	198	169	..	..
T4 wild × T2 X	228	189	..	..

* Here defined as (*total yield—unadsorbed phage*)/*bacterial input*.

In lines 1 to 6 the burst sizes are a measure of the amount of complementation between gene pairs and complementation groups indicated. In all tests identical dilutions and platings were made, so that the smaller values in the table have less statistical reliability than the larger ones. The value of "3" recorded in Experiment 2 for the test T2B×T438 cannot be considered indicative of complementation; it probably represents an inadequacy in the correction for unadsorbed phage applied to each of the yields. Expts. 3 and 4 were performed in a different connection; data from them are included as a demonstration of the reproducibility of the burst size measurements.

DISCUSSION

Much has been written about clusters of "related" genes. Almost invariably a particular discussion has been restricted to either a functional explanation (e.g. operons) or a populational explanation (e.g. supergenes). The contrast between these two hypotheses (which are *not* mutually exclusive) exemplifies the contrast between a "functional" and an "evolutionary" biologist. The functional hypothesis to explain gene clusters is a good scientific hypothesis; for *any particular case* a simple experiment has the possibility of showing it to be wrong (e.g., see BAUMBERG, BACON and VOGEL 1965). The populational hypothesis, on the other hand, is at the edge of scientific respectability; for any particular case only a complete description of the anatomy, physiology, embryology, ecology, population mating structure, and evolutionary history (and more?) could demonstrate the inadequacy of that explanation. This is a hopeless task. Since the populational hypothesis cannot be scientifically tested, are we nevertheless permitted to use it? Certainly. In fact its value must be judged through its usefulness. In the discussion which follows the orthotopic linkage system hypothesis is put to use to provide an explanation for the widespread occurrence of circular linkage maps in microorganisms. Of course, this explanation enjoys the same immunity as does the hypothesis which engenders it.

Let us rank the intensity of pairwise interaction between genes (i.e., between gene products) as 1st order, 2nd order, etc. In an orthotopic system, the recombinant frequency will be correlated with the interaction order, 1st order interacting genes being most closely linked. In the general case, many pairwise combinations of N genes may share a common interaction order. In a linear linkage system the optimal gene arrangement must often be a "compromise" arrangement since two genes cannot occupy the same locus (definition). Compro-

mises would be reduced in multidimensional linkage systems—i.e., gene arrangements could be made more nearly optimal. Therefore the one-dimensionality of existing linkage systems must reflect considerations which have restricted the evolution of optimal orthotopic systems. Perhaps multidimensional template processes are as hard to achieve as they are to conceive. Or perhaps evolution has been stuck with nucleic acid, a one-dimensional molecule, because no other will work. In any case, since nature has been restricted to one-dimensional linkage systems, we shall gain in simplicity with no relevant loss of generality if we restrict the remainder of our discussion to such systems.

A brief consideration reveals that the *elimination of ends* from linear linkage maps makes possible more nearly optimal linkage arrangements. In a map without ends *every* gene has the possibility of being equidistant from *two* genes with which it has a common order of interaction. In an ordinary linear map, the ends force worse gene arrangements. Since the removal of ends results in the removal of end effects, genetic circularity is seen to be a corollary of gene clustering.

May we emphasize that our argument refers to the circularity of the genetic *map*. This map circularity can in principle be achieved in a number of different ways, from which we have found it convenient (STAHL 1964) to distinguish three categories. (1) The interacting chromosomes are circular. (2) The interacting chromosomes are linear but have gene sequences which are circular permutations of each other. (3) The interacting chromosomes are linear with identical gene sequences, but they always engage in an even number of crossovers. Case (1) is illustrated by *Escherichia coli* (see JACOB and WOLLMAN 1961, and CAIRNS 1963). Case (2) is illustrated by T4 (see STAHL 1964) and quite possibly by Streptomyces (HOPWOOD 1965 and personal communication). Case (3) has not been reported. The discovery of case-(3) circularity, resulting from mere adjustment of the values of the classical interference parameters, would provide the most suggestive illustration of the adaptive value of map circularity *per se*.

The notions and experimental approach described herein grew out of an ambulatory discussion in London between R. H. EPSTEIN and FWS. We are grateful for the tolerance shown us by many members of the Division of Molecular Genetics of the M.R.C. Unit for Molecular Biology during the course of this somewhat frivolous study. We are especially indebted to MR. RICHARD RUSSELL of the Biology Division, California Institute of Technology, for sharing with us the unpublished results of his many experiments involving T2 and T4 amber mutants. The primary result reported in this paper, the failure of intertype complementation with genes 37 and 38, was observed by MR. RUSSELL prior to our own observation.

SUMMARY

Complementation tests between T2 and T4 amber mutants have revealed a case of a closely linked pair of polymorphic genes whose products interact in determining viability. This case and others in the literature suggest that gene clustering in microorganisms is a result in part of selection against recombinants arising during genetic exchange between quasi-wild-type individuals. If such a selective force is indeed of significant magnitude, it may account for the widespread occurrence of circular linkage maps among microbes.

LITERATURE CITED

ADAMS, M. H., 1959 *Bacteriophages*. Interscience, New York.

BAUMBERG, S., D. F. BACON, and H. J. VOGEL, 1965 Individually repressible enzymes specified by clustered genes of arginine synthesis. Proc. Natl. Acad. Sci. U.S. **53**: 1029–1032.

BODMER, W. F., and P. A. PARSONS, 1962 Linkage and recombination in evolution. Advan. Genet. **11**: 1–100.

CAIRNS, J., 1963 The bacterial chromosome and its manner of replication as seen by autoradiography. J. Mol. Biol. **6**: 208–213.

DEMEREC, M., 1965 Synthesis. Proc. 11th Intern. Congr. Genet. **2**: 51–54.

EDGAR, R. S., G. H. DENHARDT, and R. H. EPSTEIN, 1964 A comparative genetic study of conditional lethal mutations of bacteriophage T4D. Genetics **49**: 635–648.

EDGAR, R. S., and R. H. EPSTEIN, 1965a Conditional lethal mutations in bacteriophage T4. Proc. 11th Intern. Congr. Genet. **2**: 1–16. —— 1965b The genetics of a bacterial virus. Sci. Am. **212**: 71–78.

EPSTEIN, R. H., A. BOLLE, C. M. STEINBERG, E. KELLENBERGER, E. BOY DE LA TOUR, R. CHEVALLEY, R. S. EDGAR, M. SUSMAN, G. H. DENHARDT, and A. LIELAUSIS, 1963 Physiological studies of conditional lethal mutations of bacteriophage T4D. Cold Spring Harbor Symp. Quant. Biol. **28**: 375–394.

FISHER, R. A., 1930 *The Genetical Theory of Natural Selection*. Clarendon Press, Oxford.

HOPWOOD, D. A., 1965 A circular linkage map in the actinomycete *Streptomyces coelicolor*. J. Mol. Biol. **12**: 514–516.

JACOB, F., and J. MONOD, 1961 Genetic regulatory mechanisms in the synthesis of proteins. J. Mol. Biol. **3**: 318–356.

JACOB, F., and E. L. WOLLMAN, 1961 *Sexuality and the Genetics of Bacteria*. Academic Press, New York.

LEWONTIN, R. C., 1965 The role of linkage in natural selection. Proc. 11th Intern. Congr. Genet. **3**: 517–525.

SARABHAI, A. S., A. O. W. STRETTON, S. BRENNER, and A. BOLLE, 1964 Co-linearity of the gene with the polypeptide chain. Nature **201**: 13–17.

STAHL, F. W., 1964 *The Mechanics of Inheritance*. Prentice-Hall, Englewood Cliffs, New Jersey.

THOMAS, R., 1964 On the structure of the genetic segment controlling immunity in temperate bacteriophages. J. Mol. Biol. **8**: 247–253.