*DISTANCES SEPARATING GENETIC MARKERS IN T4 DNA**

By Gisela Mosig

CARNEGIE INSTITUTION OF WASHINGTON, GENETICS RESEARCH UNIT, COLD SPRING HARBOR, NEW YORK, AND DEPARTMENT OF MOLECULAR BIOLOGY, VANDERBILT UNIVERSITY, NASHVILLE, TENNESSEE

Communicated by A. D. Hershey, July 19, 1966

Populations of bacteriophage T4 include a small fraction of light particles, each containing a single DNA fragment of fixed length, measuring only two thirds of the normal molecular length.[1,2] These particles are individually nonviable, but they can cooperate to produce normal progeny when two or more of them infect a single bacterium.[2] Evidently different light particles lack different parts of the genome.

The nucleotide sequences of complete T4 DNA molecules are circularly permuted.[3,4] The following experiments show that the DNA fragments of the light particles represent continuous segments cut at random from the circular genome. Since the length of the fragments is known, the frequency with which they accommodate marker pairs is a measure of the physical distance between the markers.

Materials and Methods.—*Escherichia coli* strains B, CR63, K(λh) no. 3, and S/6 were obtained from A. H. Doermann. Most amber (*am*) mutants of bacteriophage T4D were isolated by R. H. Epstein and obtained from R. S. Edgar. The am^+ osmotic shock-resistant mutants, the r_{II} mutants, and some of the *am* mutants were isolated by A. H. Doermann and obtained from him or from F. C. Womack.

General procedures and media were described by Chase and Doermann.[5] The *am* stocks were grown in *E. coli* CR63.

Light particles were isolated by virtue of their low buoyant density. Osmotic shock-resistant mutants of phage T4D were grown in *E. coli* B in a mineral-glucose medium[6] containing P^{32}-orthophosphate (0.1–0.5 c/gm P). The lysates were clarified by low-speed centrifugation, digested with DNase and RNase, and filtered through membrane filters. The phage particles were then sedimented and resuspended in a 0.1 *M* Tris buffer (pH 7.4) containing 0.01 per cent gelatin, 0.1 per cent NaCl, and 0.2 per cent $MgCl_2$. The phage suspension was brought to density 1.5 gm/ml with cesium chloride (Harshaw, optical grade) and heated at 45°C for 30 min. Three-ml samples were centrifuged for 24 hr at 27,500 or 20,000 rpm at 20°C in a Spinco type SW39 rotor. Fractions were collected through the bottom of the tube, and analyzed for radiophosphorus and for viable particles. The fractions containing light particles were contaminated with viable particles. They were centrifuged again to obtain well-resolved fractions of light and of normal particles. The isolated light particles contained less than 1% of viable particles. The titer of defective particles was taken to be $^3/_2$ times the plaque titer of normal particles of equal P^{32} content isolated from the same tube. Whenever it was checked, the titer of defective particles, so calculated, agreed with their bacteria-killing titer.

Results.—Since two or more defective particles can cooperate to produce complete progeny, it was assumed that different defective particles lack different genes.[2] If a fixed fraction of the genome is omitted at random, all genes should be represented in defective particles with equal frequency. That frequency should be two thirds if genetic information is proportional to the DNA content of the particles.

Downloaded by guest on June 16, 2021

These expectations were tested by measuring the fraction of defective particles containing the wild-type alleles of different *am* mutations. The *am* mutants are by definition unable to grow in *E. coli* B or S/6.[7] However, they can grow in bacteria which also receive a functioning am^+ gene by simultaneous infection with a light particle. Thus, the fraction of phage yielders among bacteria simultaneously infected with a few complete *am* particles and a single light am^+ particle provides a minimum estimate of the fraction of defective particles containing the am^+ gene.

E. coli B were infected with 2–5 normal particles carrying a single *am* mutation and an average of <0.1 light am^+ particles per bacterium. Before the end of the latent period the infected cells were plated on *E. coli* S/6. In order to produce plaques on S/6, the am^+ allele has to function and to appear in at least one viable progeny particle. If both processes are perfectly efficient, two thirds of the bacteria infected with a defective particle should yield phage progeny. Results are therefore expressed as the percentage:

$$\frac{[\text{bacteria-yielding phage}]}{[\text{bacteria infected with light particles}]} \times \frac{3}{2} \times 100.$$

The main error in this measurement results from uncertainty of the quantity in the denominator. That quantity depends on the estimated titer of isolated light particles and on the efficiency of adsorption, which was assumed to be equal to the efficiency of viable particles in control experiments.

In two additional experiments bacteria were simultaneously infected with several complete r_{II} particles and a single light r^+ particle and plated on K(λ). Since only r^+ phage grow on K(λ),[8] these experiments measured the fraction of light particles containing the r_{II} gene.

The results for 19 *am* and 2 r_{II} mutants are summarized in Table 1. They show that any of the tested genes, independently of their map position, are present and function in two thirds of the defective particles. Experiments with *am* mutants were repeated as described above, except that CR63 was used instead of strain B. In CR63 the am^+ function is not required for phage replication. The infected bacteria were plated on S/6, in which only am^+ phage can grow. In this way one measures the fraction of defective particles that contribute their am^+ gene to at least one viable progeny particle. Results are included in Table 1. As in the experiments with strain B, the frequency with which different am^+ genes are recovered does not depend on their map position. This frequency proves to be lower on the average in CR63 than in B, for reasons not yet understood. Perhaps genetic recombination is not as efficient in strain CR63 as in strain B.

It is reasonable to assume that each DNA fragment represents a continuous segment of the T4 map. If so, the frequency with which a light particle can contribute two markers in mixed infection with normal particles should depend on the physical distance between the markers on the DNA molecule. This principle is used in the following experiments to measure the distance between marker pairs. *E. coli* B were infected with an average of one complete r^+ *am* particle and an average of 0.03 defective particles (genotype r_{73} am^+) per bacterium. The infected bacteria were plated on S/6, in which only r^+ am^+ and r am^+ phage can grow. The resulting plaques were classified as r^+ or mixed r and r^+ by inspection. To

Downloaded by guest on June 16, 2021

TABLE 1

EFFICIENCY WITH WHICH am^+ GENES OF LIGHT PARTICLES FUNCTION AND ARE TRANSMITTED TO PROGENY IN BACTERIA SIMULTANEOUSLY INFECTED WITH *am* MUTANTS

Gene	Mutant	Recovery of Marker as Per Cent of Expected Frequency (in B)	(in CR63)
5	135	95	76
	"		77
5	256	127	69
	"	113	
6	251	106	66
7	16	72	60
	"	97	
25	s-52	100	78
26	N131	74	45
26	S105	170	60
	"	129	
51	s-29	125	76
			64
27	s-92	76	53
27	s-60	115	55
29	s-71	112	58
29	6	84	41
	"	123	
	"	117	
48	85	117	64
	"	114	
48	s-6	92	41
31	54	108	51
34	58	96	49
37	52	76	36
	"	123	
r_{II} B	r_{73}	147	
	"	101	
r_{II} A	r_{64}	130	
41	81	94	44
	"	104	
	"	100	
44	82	72	36
46	130	130	55
	"	120	
	"	110	

The number of bacteria-yielding phage is expressed as a percentage of two thirds of the number receiving a light particle. All experiments were done with the same preparation of light particles, except for the third repetitions in *E. coli* B with the mutants 6, 81, and 130. The data are not corrected for the fraction of bacteria receiving two or more particles (multiplicity of infection <0.1).

justify this classification, many plaques scored as r^+ were sampled and checked for r particles. None were found.

Theoretically, the relative frequency of r^+ and mixed plaques generated under the conditions stated should be related to the distance between the markers as follows. Only two thirds of the defective particles contain the am^+ gene and can initiate phage production in the mixedly infected bacteria. When the distance (D) between the two markers is less than $^1/_3$ of the total molecular length, a DNA fragment containing the am^+ gene can include r_{73} only by spanning the shorter of the two distances on the circular map (see Fig. 1). Then, a fraction $P = {}^2/_3 - D$ of the defective particles contains both am^+ and r^+ genes. The fraction of productive bacteria yielding both r and r^+ phage is therefore $\frac{{}^2/_3 - D}{{}^2/_3} = \frac{3P}{2}$ and the fraction yielding r^+ only is $Y = 1 - 3P/2 = 3D/2$. The measure of distance is thus

$$D = 2Y/3. \qquad (1)$$

Downloaded by guest on June 16, 2021

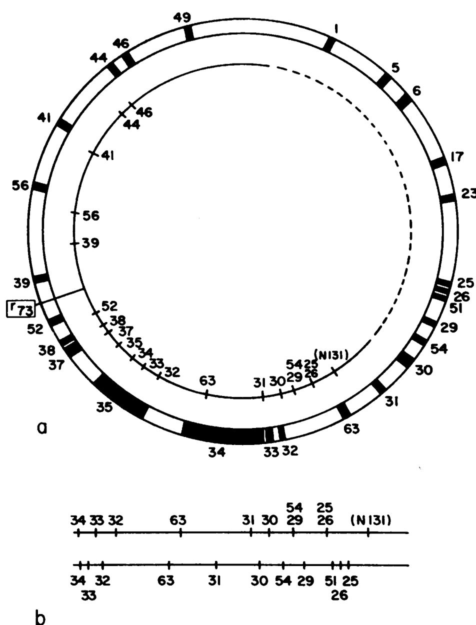

FIG. 1.—Genetic and physical maps of the T4 chromosome. The outer circle of (*a*) shows gene numbers and their positions as reported by Edgar and Wood.[9] The inner circle shows gene numbers and positions obtained by the methods described in this paper. The discrepancy between the lengths of genes 34 and 35 on the two circles distorts the comparison of all distances from r_{73} beyond gene 34. Therefore, in (*b*) the distances measured in the region between genes 34 and 25 by genetic crosses (*lower line*) are compared to the corresponding distances measured here (*upper line*).

When the distance between the *am* and the *r* mutation is greater than $^1/_3$, a DNA fragment containing the am^+ gene can include r_{73} as well by spanning either of two distances. In this case,

$$P = (^2/_3 - D) + [^2/_3 - (1 - D)] = {^1/_3}$$
$$Y = 1 - 3P/2 = 0.5.$$

Thus, all defective particles containing am^+ markers separated by more than one third of the genome from r_{73} are expected to yield half r^+ and half-mixed plaques under the given experimental conditions.

To provide the am^+ function, a DNA fragment probably has to include the whole gene. Therefore, the distance calculated by equation (1) is the distance between the r_{73} mutation and the distal end of the selected am^+ gene. Thus all mutants of the same am^+ gene should give the same results.

Distances so calculated between r_{73} and 29 different *am* mutants are summarized in Table 2. The following facts should be noted. Distances range from 0.027 to one third of the total map. As expected, the fraction of r^+ plaques never exceeds one half. This justifies the assumption, made above, that the DNA fragment represents a continuous segment of the genetic map. Mutations at the proximal and distal ends of gene 34 (A455 and B24) give identical results. The double mutant B262/52 gives the results expected for 52 alone: only the mutation distal to r_{73} determines the frequency of r^+ plaques.

Downloaded by guest on June 16, 2021

TABLE 2

DISTANCES BETWEEN THE r_{73} MUTATION AND THE DISTAL ENDS OF DIFFERENT GENES

Gene	*am* Mutant	No. of plaques	No. of r^+ plaques	Distance from r_{73}	Gene	*am* Mutant	No. of plaques	No. of r^+ plaques	Distance from r_{73}
52	H17	847	30		39	N116	821	48	
		652	31	0.027			327	26	
							576	37	0.043
38	B262	808	49						
		1118	65	0.040	56	E51	801	92	
							209	15	0.071
38/37	B262/52	734	50						
		134	12	0.048	41	81	401	88	
							1062	198	0.130
35	B252	455	46						
		621	59	0.065	44	82	1135	287	
							1061	294	
34	A455	781	94				596	158	0.177
		157	18						
	B25	339	42		46	130	552	150	
		337	46	0.081			507	148	0.189
33	N134	1405	198	0.094					
32	A453	389	68						
		1061	175	0.112					
63	E1072	1353	309						
		1387	349	0.160					
31	N54	1423	460	0.216					
30	H39	1232	430		1	B24	1150	562	
		594	198	0.230			716	342	>0.333
54	H21	1327	489		5	B256	229	118	
		995	367			135	393	196	>0.333
29	6	1248	447						
		358	142		6	251	1226	595	>0.333
	S71	410	150	0.245					
					17	N56	1037	515	
							744	375	>0.333
51?	S60	713	276						
		580	226		23	17	897	461	
							698	332	>0.333
26	S105	488	197						
		1067	432						
25?	S52	513	210						
		267	99						
		204	74	0.264					
26	N131	617	272						
		912	374						
		787	358	0.288					

The numbers of plaques shown in the table have been corrected for about 1% contamination of defective particles with normal particles. The measurements are not corrected for the redundancy of the normal T4 DNA molecule. There are two minor inconsistencies in the data, which are not yet understood. Mutants in genes 29 and 54 give identical fractions of pure r^+ plaques. The same is true for mutants in genes 25, 26, and 51. Therefore, only one distance is given for each of the two gene clusters. The mutant N131 in gene 26 gives a higher fraction of pure r^+ plaques than do three other mutants in this gene cluster. Perhaps N131 is a double mutant.

Figure 1 shows that the sequence of genes in the DNA molecule resembles the sequence of genes on the genetic map.[9] In most areas, distances on the two maps are in good agreement. There are discrepancies between the distances obtained by genetic and physical measurements, however. One is seen in the region including genes 34 and 35, a second, less convincing, in the region including genes 39 and 56.

Discussion.—The data presented above show that all genes have an equal chance to be represented in the DNA fragments of the light particles. In mixed infection of *E. coli* B with normal particles, the light particles can transmit to progeny any gene which is represented. When the frequency with which light particles contribute two markers is used as a measure of distances between different marker

Downloaded by guest on June 16, 2021

pairs, and a map is constructed from the data (Fig. 1), the resulting marker sequence agrees with the marker sequence on the genetic map.[9] These findings justify the assumption that the DNA fragments in light particles represent continuous segments of the genetic map that are cut at random from an effectively circular genome. Thus, they support the concept of circular permutation,[3] and the idea that the lengths of T4 DNA molecules are determined by a mechanism that cuts to measure. That mechanism must determine at least two distinctive lengths, the normal molecular length and the two-thirds length found in light particles. A similar bimodal distribution of lengths of DNA molecules has been found in phage P1.[10]

In most areas of the map, distances measured by the method described in this paper agree remarkably well with distances obtained by genetic crosses. In the region including genes 34 and 35, however, genetic map distances are larger than distances reported here. The discrepancies suggest that genetic recombination frequency in this region is increased by local factors other than distances between markers. Alternatively, the distance measurements described here could be in error if the ends of the DNA fragments were cut less frequently in the region mentioned than in other regions.

The second explanation appears less likely, for the following reason. Distances measured by applying equation (1) are based on the fraction of DNA fragments carrying the am^+ marker which have lost the *r* marker. Thus, they reflect the frequency of left ends of am^+-containing fragments, when the *am* marker lies within one third of the map to the right of the *r* marker (if the clockwise direction on the map is defined as going from right to left). Since the DNA fragments are uniform in size,[2] each left end must be correlated with a right end at distance two thirds of the genome. If left ends of fragments containing the genes to the right of gene 37 occurred so infrequently in the region including genes 34 and 35 as to cause the observed discrepancy, right ends of those fragments should occur with similar low frequency in the region between genes 46 and 56. These genes lie within one third of the map to the left of the *r* marker, and distance measurements in that region depend on the frequency of right ends of am^+-containing fragments. Therefore, one should expect another discrepancy between the two maps in the region including genes 46 and 56. Contrary to this expectation, the two maps agree very well in that region. It is possible that the measurements are in error at distances nearly one third from the *r* marker. Experiments with different *r* mutations as reference markers are being done to check this possibility and to provide distance measurements for that one third of the genome which gave 50 per cent pure r^+ plaques. From the data available, it appears more likely that in bacteriophage T4 there are regions in which recombination frequencies per unit length DNA are greatly increased. This would be a local effect opposite to the one reported by Jordan and Meselson[11] for bacteriophage λ. They concluded from the distribution of parental DNA in recombinant progeny particles that there is a considerable length of DNA of that phage within which genetic recombination is rare.

Two other findings might be related to the discrepancy between physical and genetic distances in T4 reported here.

The wild-type alleles of different *am* mutations are rescued from irradiated phage particles with unequal efficiency. Womack[12] found four clusters of T4 genes for which the efficiency is high. One of these includes genes 34 and 35, which are also

Downloaded by guest on June 16, 2021

characterized by high recombination frequency as shown above. In two other clusters with high efficiency of marker rescue, however, distances on the two maps agree.

Bacteriophages T2 and T4 are closely related and can exchange genetic markers by recombination. Amber mutants in T2 can be correlated with equivalent genes in T4 by complementation tests. R. Russell[13] has measured map distances between *am* mutants in T2. The genetic map of T2 differs from that of T4 and resembles the physical map of T4 given here. Thus, it appears that the region of the genome which shows high recombination frequency or low cutting frequency in T4 does not show either in T2.

It is tempting to speculate about the nature of local factors that influence recombination frequencies in T4. The DNA molecules of T2 and T4 have similar or identical phosphorus contents.[1] They sediment at very similar rates through sucrose gradients,[14] and the difference could reflect the unequal glucosylation of hydroxymethylcytosine[15] rather than unequal nucleotide contents. Perhaps there is a concentration of hydroxymethylcytosine in genes 34 and 35 whose glucosylation affects recombination frequency.

Summary.—A class of light T4 particles contains DNA fragments two thirds the length of normal T4 DNA. The fragments represent continuous random segments of the circular genome. The frequency with which the fragments contribute marker pairs in mixed infection was used to measure distances between markers. The results confirm the colinearity of the genetic map and the DNA molecule in T4. In most areas, measured distances in DNA are proportional to map distances determined from recombination frequencies. However, the region including genes 34 and 35 seems to be especially prone to genetic recombination.

It is a pleasure to thank Dr. A. D. Hershey for his encouraging interest and advice and for editing the manuscript, Drs. R. S. Edgar, A. H. Doermann, and F. C. Womack for mutant phage strains and information about them, and Drs. R. Ehring and E. Goldberg for stimulating discussions.

* This investigation was supported by U.S. Public Health Service research grants HD 01228 and GM 13221 from the National Institutes of Health.

[1] Rubenstein, I., C. A. Thomas, Jr., and A. D. Hershey, these PROCEEDINGS, **47,** 1113 (1961).

[2] Mosig, G., in *Cold Spring Harbor Symposia on Quantitative Biology*, vol. 28 (1963), p. 35.

[3] Streisinger, G., R. S. Edgar, and G. H. Denhardt, these PROCEEDINGS, **51,** 775 (1964).

[4] Thomas, C. A., Jr., *J. Gen. Physiol.*, **49,** no. 6, part 2, 143 (1966).

[5] Chase, M., and A. H. Doermann, *Genetics*, **43,** 332 (1958).

[6] Hershey, A. D., and N. E. Melechen, *Virology*, **3,** 207 (1957).

[7] 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, in *Cold Spring Harbor Symposia on Quantitative Biology*, vol. 28 (1963), p. 375.

[8] Benzer, S., in *The Chemical Basis of Heredity*, ed. W. D. McElroy and B. Glass (Baltimore: The Johns Hopkins Press, 1957), p. 70.

[9] Edgar, R. S., and W. B. Wood, these PROCEEDINGS, **55,** 498 (1966).

[10] Ikeda, H., and J. Tomizawa, *J. Mol. Biol.*, **14,** 120 (1965).

[11] Jordan, E., and M. Meselson, *Genetics*, **51,** 77 (1965).

[12] Womack, F. C., *Virology*, **26,** 785 (1965).

[13] Russell, R., personal communication.

[14] Mosig, G., unpublished experiment.

[15] Sinsheimer, R. L., these PROCEEDINGS, **42,** 502 (1956).

Downloaded by guest on June 16, 2021