A stable duplication as an intermediate in the selection of deletion mutants of phage T4

By C. van de VATE* and N. SYMONDS

*School of Biological Sciences, University of Sussex,
Brighton, Sussex*

SUMMARY

A selective procedure is described for the isolation of deletions in phage T4. This is based on the properties of partial diploids which possess a tandem duplication covering the rII region. The starting-point of the procedure is a particular s1231/3157 diploid which has a high segregation frequency and so possesses a long duplication. The replacement in this diploid of the large deletion 3157 by the small one 196 further increases the length of this duplication and removes the terminal redundancy of the resulting phage which are then non-viable. New compensating deletions which restore the terminal redundancy and thus the viability to these diploids are thereby selected. Nine new independent T4 deletions have been isolated by this procedure. The length of three has been estimated by terminal redundancy measurements and each found to be appreciably longer than the rII region. While developing the isolation procedure a key diploid intermediate was identified with the property that is was completely stable; that is, unlike typical diploids which continually generate haploid segregants during phage multiplications, this diploid strain produced no viable haploid segregants. This unexpected finding led us to examine the general problem of how tandem duplications can be stabilized. The solution we propose here is a structural one involving new stabilizing deletions which penetrate into one arm of the duplication removing certain essential genes. These stabilizing deletions are located in such a way that all haploid segregants formed by recombination are necessarily non-viable. Further investigations of the original stable diploid, and others isolated subsequently, validated this model for these phage diploids. It also led to the recognition of several novel genetic structures involving repeated DNA sequences which, together with the concept of the stabilizing deletions, could be of general significance.

1. INTRODUCTION

The genetic map of phage T4 shows many regions not covered by essential genes (Edgar & Wood, 1966). Over the last few years an increasing number of non-essential genes have been identified and mapped (e.g. Tessman & Hall, 1967; Warner, Snustad, Jorgensen & Koerner, 1970; Warner, Snustad, Koerner &

Childs, 1972). In order to clarify the roles of these genes, and also to define the extent of the non-essential regions in the T4 genome, a method of generating T4 phage deletions is necessary.
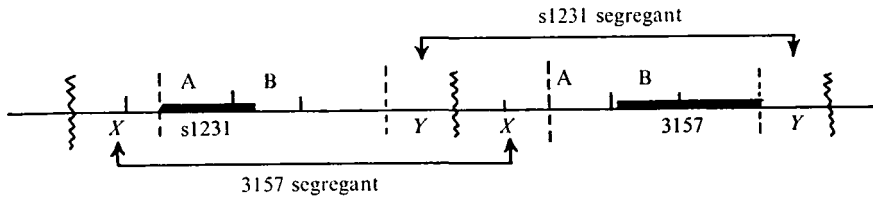

Fig. 1. Structure of an rII diploid of the type s1231/3157. The squiggly lines represent the limits of the duplication: X and Y the extent of the two homologous regions within the duplication. A recombinational event in the X region between the two parts of the duplication leads to the formation of a haploid 3157 segregant; in the Y region to a haploid s1231 segregant.

In this paper we describe a method based on the properties of the rII diploids of phage T4 (Weil, Terzaghi & Crasemann, 1965; Symonds *et al.* 1972; Parma, Ingraham & Snyder, 1972). These rII diploids arise from crosses between an rII A deletion such as s1231 (which penetrates into the B cistron (see Fig. 2) but still retains B functional activity; Drake, 1963) and a deletion like 3157 which overlaps s1231 but is entirely in the B cistron and so complements s1231 because it has A functional activity. No wild-type recombinants can arise from such a cross but phages are formed at low frequency (the so-called rII diploids) which can multiply in bacteria lysogenic for phage λ. These phage contain a tandem duplication, one part of which is inherited from each parent in the cross. The structure of such a duplication for a s1231/3157 diploid is shown in Fig. 1. Two regions of homology exist in the structure, labelled X and Y in the figure. Recombination events between these regions lead to the formation of haploid segregants. The frequency with which these segregants occur during a single cycle of phage growth (the segregation frequency) is characteristic for a particular rII diploid and is determined by the overall length of the duplication. The ratio of s1231 to 3157 segregants (the segregation ratio) depends on the relative sizes of X and Y.

An important implication of the structure is that rII diploids have a smaller terminal redundancy than wild-type T4. As the duplications become longer, the terminal redundancy must eventually disappear. It would be expected that phages harbouring such long duplications would be non-viable because it is generally considered that the terminal redundancy plays some essential role in phage DNA synthesis. A way of restoring terminal redundancy and therefore viability to diploids of this type is to introduce an extra deletion in their genome. One method of looking for these compensating deletions is to assume that such deletions are already present in some of the spontaneously occurring rII diploids. Weil & Terzaghi (1970) have successfully isolated and mapped new deletions obtained in this way. Another method is to start with a particular diploid, say

of the type s1231/3157, and then substitute in it the much smaller deletion 196 for 3157. This substitution (see Fig. 2) will increase the length of the duplication by over 1500 base pairs. With s1231/3157 diploids for which the segregation frequency is high (that is for which the extent of the duplication is already large) it might be expected that this further increase in the duplication would lead to non-viability of the s1231/196 phage unless a compensating deletion was simultaneously introduced into the phage genome to restore the terminal redundancy. This rationale allows the isolation of deletions and may permit a systematic investigation of the non-essential regions of the T4 genome.

During these experiments a class of *stable* rII diploids was found. The interest of this finding is that one would expect that all tandem duplications should be inherently unstable because they continually generate haploid segregants as they reproduce. This is normally true of the rII diploids, and there are many other examples of this behaviour in micro-organisms (for example the instability of recombinants formed by specialized transduction), and also numerous instances have been reported with eukaryotes. However, particularly in eukaryotic systems, it is common to find repeated DNA sequences in chromosomes, and it has been supposed that gene duplication plays a major role in evolution (Ohno, 1970). If this is a correct interpretation, then some means of stabilizing duplications must have been developed at an early stage in evolution. We were, therefore, led to examine possible ways in which this stabilization could occur.

2. MATERIALS AND METHODS

(a) *Bacteria*. The origin of the various strains is described in Symonds *et al*. (1972). For convenience, permissive hosts on which rII mutants will grow are referred to as B, and non-permissive hosts as K.

(b) *Phage*. The map locations of the T4 rII mutants used in this study are indicated in Fig. 2. In the deletion mutant 1231, as in 1589, a non-essential part of the B gene is deleted. However, the B gene is not functional in the case of 1231 because the reading frame is altered. The point-mutant s12 is a frame-shift mutation such that the double-mutant s12–1231 (here abbreviated to s1231) has B functional activity and behaves in a similar way to 1589.

(c) *General techniques*. The experimental conditions used for the isolation of rII diploids, spot-tests, estimates of terminal redundancy and the measurement of segregation frequencies and ratios are all similar to those described previously (Symonds *et al*. 1972).

(d) *Nomenclature*. We shall refer to the rII partial diploids simply as rII diploids. These contain a tandem duplication having at its centre the repeat point designated by the sign /. The designation s1231/3157 therefore refers to a heterozygous diploid containing different genetic information (s1231 and 3157) in the two parts of the duplication. Homozygous diploids will be referred to in the same way. These designations do not imply a particular orientation of the two parts of the duplication with regard to the neighbouring Ac gene. If such an

orientation is known and relevant it will be referred to as $s1231_1/3157_2$. The nomenclature of the partial rII triploids is similar to that of the diploids. The notation rII$^+$ refers to phages possessing a functional rII region.

(e) *Spot-test classification*. It is possible to distinguish six basic types of spots when rII mutants are tested against standard deletions on plates seeded with K bacteria.

(1) An empty, turbid spot indicating *no interaction* (e.g. 3157 on to W8-33). These are overlapping, non-complementing deletions so neither complementation nor recombination can lead to phage growth.
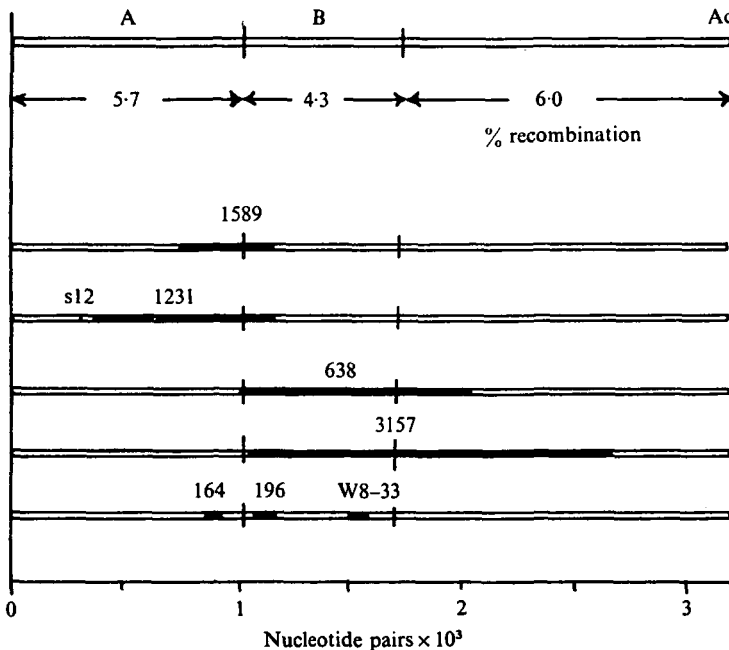

Fig. 2. Approximate map of the rII mutations used in these experiments. The extent of the deletions and the physical scale (in nucleotide pairs) is based on the results of Dove (1968).

(2) A spot with a few wild-type plaques indicating *recombination only* (e.g. 196 on to W8-33). These are non-overlapping, non-complementing deletions so a few wild-type recombinants are produced during the limited phage growth which goes on in the K bacteria due to leakiness of the rII mutants.

(3) A spot with many minute plaques indicating *complementation only* (e.g. 196 on to s1231). These are overlapping, complementing deletions and the minute plaques arise because of the complementation, but no wild-type recombinants can be formed. An occasional large plaque due to spontaneous diploid formation may be found.

(4) A fully clear spot indicating *complementation plus recombination* (e.g. s1231 on to W8-33). These are non-overlapping, complementing deletions and the

massive lysis is due to the growth of wild-type recombinants initially produced in K cells where complementation had occurred.

(5) A spot with many minute plaques but also containing 10–50 much larger plaques with an *r* morphology. These indicate *complementation plus limited recombination* and are characteristic for *segregating homozygous diploids*. Such spots occur for instance with s1231/s1231 diploids on to 3157. Because they must be grown in a permissive host any stock of s1231/s1231 only contains about 0·1 % actual diploid phage, the rest being s1231 haploid segregants. In the spot most of the phage growth is due to complementation (and gives rise to the minute plaques), but the s1231/s1231 phage also give rise to a few s1231/3157 diploids which grow as *r* plaques on the K bacteria.

(6) A spot with many minute plaques and also a larger number of *r* plaques than in (5). These also indicate *complementation plus limited recombination* and are characteristic for *stable homozygous diploids*. The number of *r* plaques in the spot reflects the frequency with which rII⁺ recombinants are formed in the relevant cross.

(*f*) *Stability tests.* The degree of stability of a suspension of lysate originating from a homozygous diploid of the type s1231/s1231 was estimated by replating the suspension on B bacteria and then spot-testing individual plaques against 3157. The fraction of plaques which gave the spot-types (5) or (6) was then used as an index of the stability of the particular homozygous diploid.

3. RESULTS

(i) *Construction of the selective system*

(a) *Experimental design*

The selective procedure for the generation of T4 deletions outlined in the Introduction depends on the isolation of an s1231/3157 diploid with a terminal redundancy which is sufficiently small that when 3157 is replaced by 196 in the diploid the terminal redundancy disappears and consequently the resulting s1231/196 phages are non-viable unless they fortuitously contain an extra deletion. Experimentally the search for such a diploid involves the following steps. (1) The isolation of a set of s1231/3157 diploids with different segregation frequencies. (2) The isolation from these of a corresponding set of s1231/s1231 homozygous diploids (*hd*). (3) A comparison of the frequencies with which s1231/196 and s1231/3157 diploids are formed as recombinants in crosses of the *hd* with 196 and 3157; the selective situation occurs when this ratio is extremely low, indicating the non-viability of most s1231/196 diploids.

(b) *Isolation of s1231/3157 diploids*

Originally 55 independent s1231/3157 diploids were isolated from the progeny of crosses between s1231 and 3157. From these a set of 5, designated A–E, was chosen with segregation frequencies ranging from 8 to 49 %. From each a sample of segregants was analysed by spot-tests. Table 1 shows that the four diploids

A–D segregate mainly 3157 phage plus a few segregants with the genotype s12–3157. It follows from the diploid structure of Fig. 1 that the only orientation of the duplication in these diploids which satisfies these segregation results is $s1231_1/3157_2$, with the 3157 deletion either extending as far as the right-hand boundary of the duplication, or very close to it.

Table 1. *Segregation properties of the s1231/3157 diploids A–E*

		Types of segregant		
Diploid	Segregation frequency (%)	s1231	3157	s12–3157
A	8	1	17	2
B	12	0	14	0
C	19	0	15	0
D	38	0	101	4
E	49	56	26	0

The diploids were plated on K and samples of these plaques replated on mixed B + K indicator. The fraction of turbid plaques was determined and the values corrected for the differential loss of segregants occurring on the indicator. The turbid plaques were then characterized by spot tests against W8–33, 196, 164 and s1231.

Table 2. *Properties of the homozygous diploids hd1–hd4*

Diploid parent	hd isolate	Recombination frequencies		Stability
		× 3157	× 196	
A	hd1	$7 \cdot 3 \times 10^{-4}$	$7 \cdot 6 \times 10^{-4}$	3/60
B	hd2	$5 \cdot 9 \times 10^{-5}$	$2 \cdot 7 \times 10^{-5}$	0/35
C	hd3	$1 \cdot 6 \times 10^{-5}$	$1 \cdot 4 \times 10^{-5}$	0/40
D	hd4	$1 \cdot 0 \times 10^{-2}$	$1 \cdot 4 \times 10^{-5}$	145/145

All crosses were performed in B using a multiplicity of approximately three of each parent. The recombinant frequency refers to the ratio of the number of progeny phage producing plaques on K bacteria as compared to those on B. To obtain the figures in the last column the *hd* isolates were plated on B and the continued presence of *hd* page in these plaques was investigated by spot tests (see Materials and Methods).

(c) Isolation of the homozygous diploids

Each of the 5 s1231/3157 diploids was crossed with s1231 and the progeny tested for the presence of *hd* of the s1231/s1231 type. Such *hd* isolates were obtained from diploids A–D, but not from E. This result probably reflects the fact that the terminal redundancy of this high-segregating diploid is already so short that the substitution of 3157 by s1231 in the duplication already leads to non-viable *hd* recombinants.

(d) Crosses with the homozygous diploids

An *hd* isolated from each of the diploids A–D (here termed *hd1–hd4*) was crossed both with 196 and 3157 and measurements made of the frequency with which rII⁺ diploids appeared in the progeny. With *hd1*, *hd2* and *hd3* the two diploids

s1231/196 and s1231/3157 are produced in comparable numbers (Table 2). However, with *hd*4 the cross against 196 yields almost 1000 times less diploids than the control cross against 3157. Therefore *hd*4 is the obvious candidate for our selective system and the few recombinants that are formed with 196 can be expected to harbour additional deletions which restore the terminal redundancy and with it the viability to these s1231/196 diploids.

The relatively high number of recombinants formed in the *hd*4 × 3157 cross is noteworthy. Homozygous diploids must be propagated in a permissive host, and due to the continuous formation of haploid segregants the proportion of *hd* in a growing culture continually decreases. The *hd* are thus inherently unstable, and the higher the segregation frequency the greater the instability. As *hd*4 is derived from diploid D, which has the highest segregation frequency of the four diploids under investigation, we would expect it to be the least stable of the four isolates *hd*1–*hd*4 and so produce the fewest recombinants with 3157. But the result with *hd*4 is contrary to this expectation. The implication is that *hd*4 is exceptional in that it has a high degree of stability.

(ii) *The unique properties of* hd4

(a) *Test for stability*

Each *hd* isolate was plated on B and individual plaques were tested to see if they retained the *hd* characteristics (see Materials and Methods). The results are shown in the last column of Table 2. With *hd*1, *hd*2 and *hd*3 reisolates that behaved as *hd* were rare; however, with *hd*4 all 145 reisolates that were tested retained the character. It appears then that *hd*4 is completely stable.

(b) *How to stabilize a duplication?*

The inherent instability of a tandem duplication is due to the continuous formation, during replication, of haploid segregants produced by recombination. Two distinct methods can be envisaged which could increase the stability of a particular duplication. The first is simply to suppress recombination. An example where this procedure has been applied with modest success is with an rII diploid of the type 1589/638 into which a partial *rec⁻* mutation was introduced leading to the reduction of segregation frequency from 25 to 7 % (Van den Ende & Symonds, 1972). In the other method segregation still occurs at the normal rate but a situation is created in which the haploid segregants are non-viable. A simple example of this is the propagation of typical heterozygous rII diploids. When *coli* B is used as a host these diploids are unstable as the rII segregants are viable and their proportion continually increases in the phage population. However, when grown in K the rII diploids are stable as the haploid segregants are then non-viable. A more general method of ensuring that haploid segregants are non-viable arises when the duplication contains genetic information which is essential for the growth of the organism. The introduction of certain types of stabilizing deletions (*sd*) into one arm of this class of duplication leads to complete stability. This is illustrated in Fig. 3.

Any deletion in the left arm of the duplication which covers part (or all) of the essential genes and runs through to the left extremity of the duplication (and probably into some non-essential region beyond) will lead to non-viability of the haploid segregants as these will inevitably have lost the one complete set of essential genes originally harboured in the other area of the duplication. The same result holds for deletions in the right arm of the duplication which cover the essential genes and run to the right-hand extremity of the duplication.
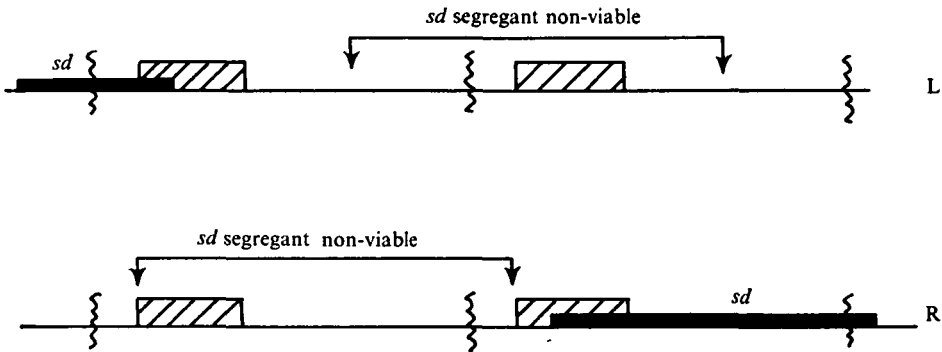

Fig. 3. Two structures for the stabilization of a duplication. Essential genes are denoted by the cross-hatched regions, the stabilizing deletion is blocked in. In both cases the haploid segregants are non-viable because they lack essential genetic information.

This structural mechanism for ensuring non-viability of the segregants due to the introduction of stabilizing deletions into duplications is of general application and we shall now explore some of its genetic consequences as applied to the rII diploids.

(c) *Models for the stability of* hd4

Although many rII deletions have been isolated which extend to the right of the B cistron in phage T4, none has been reported which extend to the left of the A cistron (Benzer, 1962). This implies there are essential genes in the latter position, probably the DNA-delay genes 39 and 60 (Sederoff, Bolle & Epstein, 1971). Most rII diploids would be expected to include part at least of this essential region within the boundaries of their duplication, so the ideas presented in the previous section are applicable to them.

In Fig. 4 four possible models are depicted which could explain the stability of *hd*4. The first assumes extra mutations affecting recombination have been introduced into the phage genome. The other three are variants of the structural mechanism. Both models II and III involve the introduction of an *sd* into the left arm of the duplication, the difference being that in model II the *sd* does not cover the rII region, while in III it does. In model IV the *sd* is in the right arm of the duplication and of necessity completely deletes the rII region there.

It is worth noting that in the L models II and III, all of the essential genes

which lie immediately to the left of the A cistron must be included within the duplication in order for them to be functional. This is not necessarily true in model IV. Also only type II of the structural models could stabilize a heterozygous rII⁺ diploid, as in this case functional copies of the rII region must still be present in both parts of the duplication. An example of the stabilization of a 1589/638 diploid by a *sd* of type II is reported in the succeeding paper (Van de Vate, Van den Ende & Symonds, 1974).
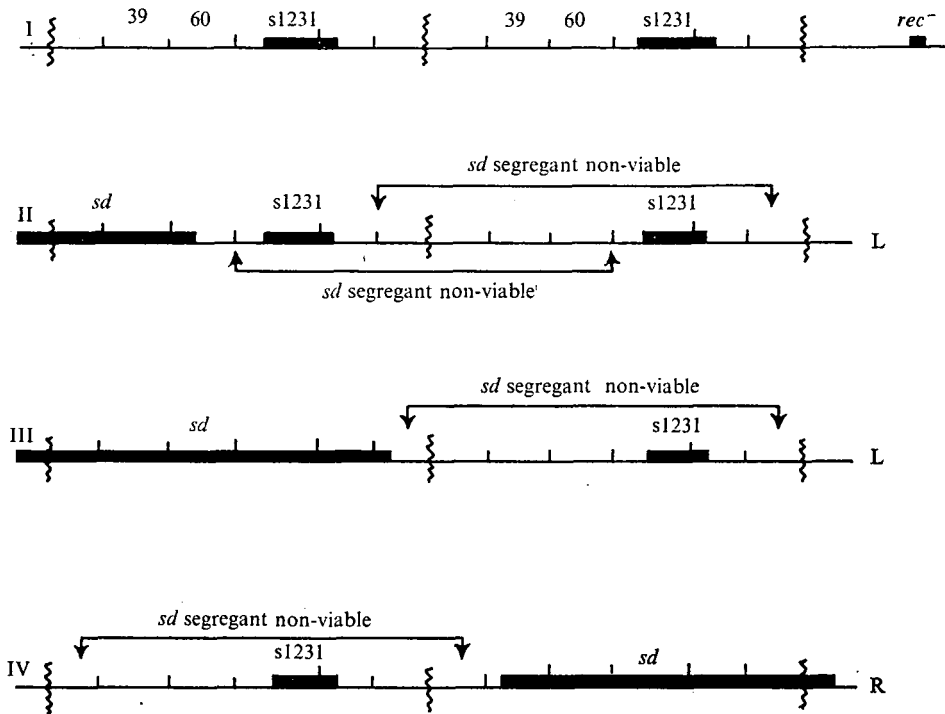

Fig. 4. Four possible models for the stable diploid *hd*4. Genes 39 and 60 are located immediately to the left of the A cistron and are essential for phage growth.

(d) *Test of the models for* hd4

The rII⁺ recombinants formed in the cross *hd*4 × 3157 were analysed. In section i (*b*) it was argued that 3157 probably extends to the right-hand boundary of the D duplication. This implies that all recombination events within the duplication must occur to the left of 3157 so that only rII⁺ recombinant diploids with the orientation s1231₁/3157₂ can arise from this cross. On model I about half of these recombinants would inherit the *rec⁻* property from the *hd*4 parent and consequently be stable. With model II at least half the recombinants would inherit *sd*, which would make them stable too. With model III it is not possible to generate rII⁺ recombinants by a single crossover, but recombination must pass through a triploid intermediate. This is illustrated in Fig. 5. First the triploid intermediate *sd*/s1231/s1231 is formed by recombination between two *sd*/s1231 genomes aligned

in a staggered position. Then recombination with 3157 will lead to sd/s1231/3157 triploids which should have the rII$^+$ phenotype. These are not stable, but would segregate stable sd/3157 segregants. A further recombination event between 3157 and the sd/s1231/3157 triploid will eventually yield s1231/3157 diploids which should be identical to the original D diploid. In model IV, s1231/3157 diploids can be recovered by a single crossover, and also will arise via the triploid intermediate.
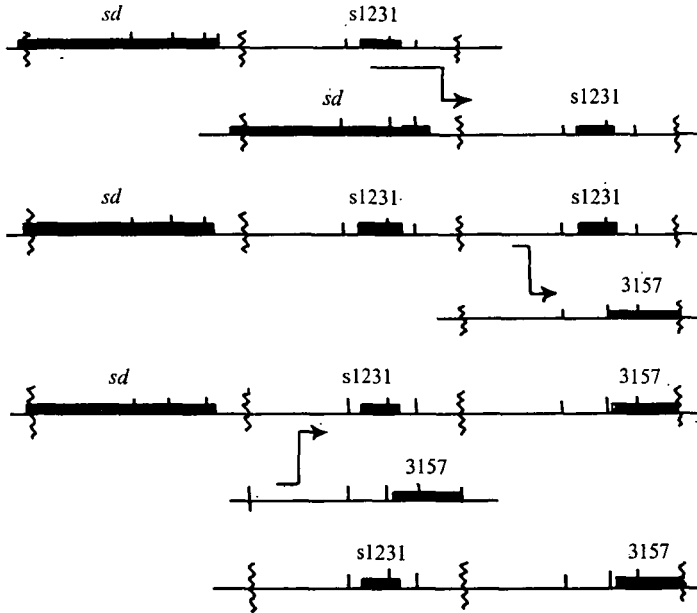

Fig. 5. Pathways of recombination by which a stable $hd4$ diploid with the structure of type III (Fig. 4) can form sd/s1231/3157 and s1231/3157 recombinants via the triploid intermediate sd/s1231/s1231.

In summary therefore models I and II predict the formation of *stable* s1231/3157 recombinants. On model III most rII$^+$ recombinants should be triploids of the type sd/s1231/3157 (yielding stable sd/3157 segregants), with a few s1231/3157 diploids. While on model IV all the recombinants should be s1231/3157 diploids which are identical to the original D diploid.

In order to test these predictions we isolated over 100 rII$^+$ recombinants from crosses between $hd4$ and 3157. None was stable s1231/3157 diploid, arguing against models I and II. To distinguish between models III and IV the segregation patterns of ten independent rII$^+$ recombinants were further investigated. They all showed segregation frequencies between 30 and 35 %, and a sample of 151 segregants showed no s1231, 5 s12-3157, and 146 3157 segregants, none of which were stable sd/3157. All ten therefore appeared to have the same segregation pattern as the original D diploid (Table 1).

Only model IV therefore satisfies the requirements for the structure of $hd4$. Clearly this type of evidence does not prove that $hd4$ has this structure; it is

conceivable that a quite different set of assumptions concerning stability could be made which lead to predictions satisfying the above results. However in the remainder of this paper we shall consider that *hd4* does have a structure of type IV and refer to the stabilizing deletion as *sd1*. (Another series of experiments pointing to the same conclusion is reported by Van de Vate, 1972.)

(iii) *Analysis of recombinants from the cross* hd*4 × 196*

(a) *Identification of deletions in the recombinants*

It has been argued earlier that any rII$^+$ recombinants arising from the cross *hd4* × 196 should contain new deletions in order to restore terminal redundancy and thus viability to the phage. Nine independent recombinants were therefore

Table 3. *Segregation characteristics of the recombinants b1–b9*

Recombinant	Segregation frequency (%)	Types of segregant	
		s1231	196
b1	42	5	19
b2	36	11	13
b3	15	2	28
b4	43	7	17
b5	39	14	31
b6	11	24	5
b7	8	9	3
b8	40	19	5
b9	35	11	13

The segregation characteristics were analysed by the methods outlined in the legend to Table 1.

isolated from crosses of this type and the presence of the extra deletions was investigated. If these recombinants contained the original D duplication with 3157 replaced by 196 then they should all display the same segregation characteristics; in particular their segregation frequency should be somewhat higher than the 35 % characteristic of D. Table 3 lists the segregation characteristics of these 9 rII$^+$ recombinants, designated b1–b9. It is clear that both the segregation frequencies and ratios vary considerably amongst the recombinants. One complication which could influence these figures is that in the cross *hd4* × 196 there are two pathways of recombination to be considered. The first is by a single crossover which yields s1231/196 diploids and are expected to resemble D. The other passes through the triploid intermediate s1231/s1231/*sd1* (formed in an analogous way to that shown in Fig. 5). A crossover with 196 then yields the triploid 196/s1231/*sd1*, which has the rII$^+$ phenotype and segregates stable 196/*sd1* and *hd4* diploid segregants, but with different segregation characteristics to D.

It became necessary therefore both to determine whether the recombinants b1–b9 possessed extra deletions, and also to find out whether they were s1231/196 diploids, 196/s1231/*sd1* triploids, or had some other structure. We proceeded to

7-2

Table 4. *Recombination frequencies obtained in crosses with the s1231 and 196 type segregants derived from the nine recombinants isolated from the cross hd4 × 196*

Recombinant	s1231-type segregants					196-type segregants				
	× 3157	× 196	× 638/sd1	Deletion	Haploid (h), diploid (d)	× s1231	× hd4	Deletion	Haploid (h), diploid (d)	Recombinant type
b1	6×10^{-5}	1×10^{-5}	2×10^{-3}	+	h	8×10^{-6}	2×10^{-3}	+	h	s1231/196
b2	1×10^{-2}	2×10^{-3}	.	+	d	1×10^{-3}	.	+	d	196/s1231/sd1
b3	7×10^{-3}	2×10^{-5}	.	−	d	4×10^{-6}	2×10^{-3}	+	h	Irregular
b4	5×10^{-5}	9×10^{-6}	2×10^{-3}	+	h	1×10^{-4}	.	+	d	Irregular
b5	6×10^{-5}	1×10^{-5}	3×10^{-3}	+	h	6×10^{-6}	2×10^{-3}	+	h	s1231/196
b6	8×10^{-3}	8×10^{-6}	.	−	d	4×10^{-6}	3×10^{-5}	−	d*	Irregular
b7	2×10^{-1}	2×10^{-1}	.	+	d	2×10^{-1}	.	+	d	Irregular
b8	1×10^{-2}	2×10^{-3}	.	+	d	1×10^{-3}	.	+	d	196/s1231/sd1
b9	4×10^{-5}	9×10^{-6}	2×10^{-3}	+	h	3×10^{-6}	1×10^{-3}	+	h	s1231/196
Controls										
s1231	3×10^{-5}	1×10^{-6}	7×10^{-6}	.	.	.	.	.	.	.
hd4	8×10^{-3}	1×10^{-5}	.	.	.	.	.	.	.	.
3157	.	.	.	.	.	2×10^{-5}	.	.	.	.
196	.	.	.	.	.	7×10^{-6}	.	.	.	.
b6	2×10^{-2}	1×10^{-2}	.	+	d	.	.	.	.	.

Crosses were performed in duplicate in *E. coli* B. The recombinant frequency is the ratio of the progeny titre on K as compared to that on B. The last line in the table refers to a second type of s1231(b6) segregant isolated from plating b6 on mixed indicator.

* This designation cannot be deduced from the results of the crosses but was arrived at after additional tests with s1231(b1).

investigate these questions by studying the properties of the segregants of these nine recombinants, as these should contain the same deletions as the parents. Consequently we isolated two segregants from each recombinant (one of each segregant type) and designated these s1231(b1), 196(b1), etc.

A decision as to whether the s1231 segregants are s1231 haploids or *hd*4 diploids can be obtained by crossing them with 3157, as nearly 300 times more recombinants should arise in the latter case (Table 4). Reference to the second column of Table 4 shows that the s1231 segregants derived from b1, b4, b5 and b9 are haploid, the others diploid. In order to determine whether the haploid segregants contained an extra deletion these were crossed to the stable diploid 638/*sd*1. The recombinant frequency between haploid s1231 and 638/*sd*1 is less than 10^{-5}, presumably because the s1231/638 diploids (like s1231/196) lack terminal redundancy. However, the frequency with the four haploid s1231 segregants is in all cases greater than 10^{-3}, strongly arguing for the presence of extra deletions in each isolate. In the case of the diploid segregants a cross with 196 then detects those which contain an extra deletion as these should show comparable recombination frequencies against 196 and 3157. From this viewpoint the diploid segregants from b2, b7 and b8 have the extra deletion, but not b3 and b6.

It is more difficult to identify the structure of the 196 segregants so accurately. As in the case of the s1231 segregants they can be of four types: 196, 196 with deletion, 196/*sd*1 and 196/*sd*1 with deletion. They were first crossed with s1231 in which case only 196/*sd*1 segregants with a deletion will yield an excess of rII$^+$ recombinants over the control cross 196 × s1231. On this basis the segregants from b2, b4, b7 and b8 are stable 196/*sd*1 diploids with an extra deletion. In order to tell whether the other segregants contained extra deletions they were crossed against *hd*4. Of these, b1, b3, b5 and b9 show recombinant frequencies at least 100 times greater than the control cross 196 × *hd*4, showing they contain extra deletions. Presumably, then, these four segregants are all haploid, as otherwise they would have yielded large numbers of recombinants in the previous cross. This then leaves just 196(b6), which from further tests appears to be a diploid segregant that does not possess an extra deletion (see legend to Table 4).

(b) *Interpretation of the deletion patterns*

The most important result arising from all these crosses is that at least one of the segregants derived from each of the recombinants appears to contain an extra deletion except in the case of b6; and upon further investigation evidence of a deletion was found there also, as two kinds of s1231 segregants could be isolated (see last line in Table 4). The philosophy of the selection system is therefore vindicated, and in each of the recombinants b1–b9 a new deletion has been isolated. We shall call these deletions *del*1–*del*9.

In order to decide on the actual structure of the nine recombinants the data in Table 4 must be considered in conjunction with that of Table 3, referring to their segregation characteristics. The three recombinants b1, b5 and b9 each generate haploid segregants containing extra deletions and display the high segregation

frequency characteristic of the D diploid. They are therefore examples of rII diploids containing the duplication s1231/196, coupled with an extra deletion which restores terminal redundancy. The recombinants b2 and b8 each generate diploid segregants containing deletions and again have high segregation frequencies so these are probably 196/s1231/*sd*1 triploids. The properties of these five recombinants are therefore just those anticipated. They contain new deletions outside the region of the duplication. The other four recombinants b3, b4, b6 and b7, however, displayed properties we did not expect and necessitated some modification of our original ideas.
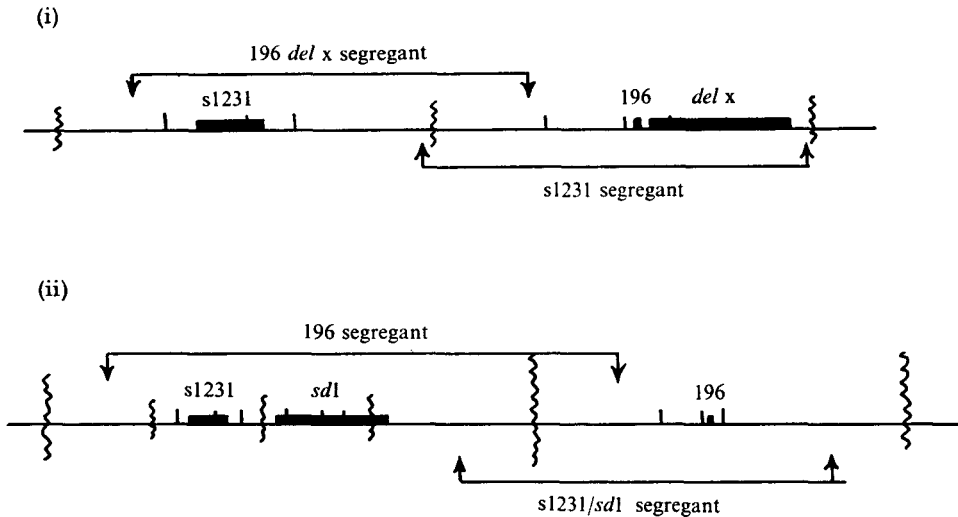
(i)

(ii)

Fig. 6. (i) Structure of a s1231/196 diploid which contains a secondary deletion, *del* x, within the right arm of the duplication. (ii) Structure of a *dudu* in which the left arm of the larger duplication contains the smaller duplication s1231/*sd*1.

(c) *The irregular recombinants*

The recombinants b3, b4, b6 and b7 show a variety of unexpected properties. Three (b3, b6 and b7) show abnormally low segregation frequencies and disturbed segregation ratios. In two cases (b3 and b6) only one of the segregants contains an extra deletion, and with two (b3 and b4) one type of segregant is diploid while the other is haploid. In addition b6 yields two kinds of s1231 segregants – one with, the other without the deletion. In this section we shall only consider in general terms the likely explanations for this behaviour; details of the structures have been reported elsewhere (Van de Vate, 1972).

The two previous structures we considered for the recombinants arising from the hd4 × 196 cross were either s1231/196 diploids or 196/s1231/*sd*1 triploids, in both cases with a compensating deletion *outside* the repeated regions in order to restore the terminal redundancy. The compensating deletion could, however, be

within the repeated regions. Such a deletion would obviously decrease the segregation frequency of the recombinant. Also according to the position of the deletion within the duplication (or triplication) either one or both of the segregants will contain it; and in certain positions two different kinds of s1231 or 196 segregant can be produced from a single recombinant. A simple illustration of this type of structure is shown in Fig. 6(i). This refers to a s1231/196 diploid in which the extra deletion (*del* x) is so positioned that s1231 and 196-*del* x segregants would be produced.

Such a deletion within the repeated structure cannot, however, explain the situations where one type of segregant is haploid and the other diploid. A possible interpretation of this behaviour is linked with the observation (Table 4) that the frequency with which recombinants arise in the cross *hd*4 × 196 is only about 10 times higher than the spontaneous frequency with which rII diploids occur by illegitimate recombinations in the cross s1231 × 196. Such illegitimate recombination events would also be expected to occur in the *hd*4 × 196 cross, and so about 10 % of the rII$^+$ recombinants could occur via the formation of a new duplicated structure in which one part of the new duplication includes the whole of the previous s1231/*sd*1 duplication. Obviously this new 'duplication of a duplication' or '*dudu*' for short must have an extensive extra deletion coupled with it. An example of this *dudu* structure is shown in Fig. 6(ii). Upon segregation this structure would yield diploid s1231/*sd*1 and haploid 196 segregants.

(d) Spot-tests and the recognition of deletions

In establishing the validity of our selection procedure we tested for the presence of deletions in the nine recombinants isolated from the cross *hd*4 × 196 by actually crossing segregants from each with a variety of reference phage. This procedure can however be speeded up tremendously by replacing the crosses with simple spot-tests and noting the pattern of response (see Materials and Methods). This means a rapid method is available both for the screening of new deletions that may arise, and also for identifying the presence of known deletions that are crossed into new genetic backgrounds. The use of the stable diploids s1231/*sd*1, 638/*sd*1 and 196/*sd*1 in these tests should facilitate the mapping and further analysis of T4 deletions which arise either spontaneously or via our selection system.

(e) Terminal redundancy measurements

So far the evidence for the new deletions is that they restore viability to the s1231/196 diploids derived from *hd*4. To obtain independent confirmation of their existence, and also to estimate their length, we measured the terminal redundancy of haploid s1231 segregants derived from recombinants b1, b2 and b4 – that is, which contained the presumed deletions *del* 1, *del* 2 and *del* 4. Similar measurements were also made with a haploid s1231 derivative obtained from *hd*4, and with our original laboratory stock of s1231.

The results in Table 5 give a clear indication of the presence of deletions in the

three recombinants and also of the presence of a small deletion in $hd4$ itself. A rough estimate of the lengths of these deletions can be made by using the relation deduced previously from similar experiments that an increase of 0·1 in the percentage of h^{2+}/h^{4+} heterozygotes represents an increase of about 375 base pairs in their terminal redundancy (Symonds *et al.* 1972). The lengths of the extra deletions calculated on this basis are also shown in Table 5. As b1, b2 and b4 were derived from crosses with $hd4$ the deletions identified here in these recombinants may or may not include the 1200 deleted base pairs inherited from this parent. Even allowing for this, however, the smallest deletion (*del* 1) is at least 2600 base pairs long, while the longest (*del* 2) is at least 7400 base pairs in extent, which is over four times the overall length of the rII region.

Table 5. *Frequency of host-range terminal redundancy heterozygotes obtained with various haploid s1231 derivations*

s1231 parent	No. of clear plaques scored	h^{2+}/h^{4+} heterozygotes (%)	Estimated lengths of extra deletions in base-pairs
s1231	31	0·52	0
s1231 (*hd4*)	52	0·85	1200
s1231 (b1)	93	1·54	3800
s1231 (b2)	177	2·85	8600
s1231 (b4)	98	1·60	4100

Progeny phage from the $h^{2+} \times h^{4+}$ crosses were plated on a mixed indicator $B + B/2 + B/4$ at about 30 plaques per plate to avoid overlaps. Fully clear plaques were attributed to h^{2+}/h^{4+} terminal redundancy heterozygotes. The phage designated simply s1231 refers to our normal laboratory stock. Haploid s1231 segregants were obtained from $hd4$ by crossing it twice with s1231 at a ratio of 1:10, and each time reisolating haploid s1231 phage by spot tests. Haploid s1231 (b2) segregants were similarly obtained from the original diploid segregant (cf. Table 4) by back-crossing with the haploid s1231 (*hd4*). The estimates of the lengths of the deletions given in the last column are obtained from the relationships derived previously that an increase in the h^{2+}/h^{4+} heterozygote frequency of 0·1 corresponds to an increase in redundancy of 375 base-pairs.

4. DISCUSSION

The aim of these experiments was to develop a selective system for isolating deletions in page T4. This has been successfully achieved. However, the key role played by the stable diploid strain $hd4$ in this isolation procedure led us to examine the question of how duplications, which are inherently unstable genetic structures, can be stabilized. This aspect of the work is potentially of general significance because of the importance of duplicated genes during the course of evolution. The mechanism by which $hd4$ and other rII diploid strains acquired stability involves the introduction of new deletions into one arm of the duplication. These stabilizing deletions cover essential phage genes and are located in such a way that the haploid segregants, which are continually being formed during phage multiplication, are inevitably non-viable. Although the indirect evidence for the existence of $sd1$ (the deletion involved in stabilizing the diploids discussed in this

paper) is overwhelming, we were not able for technical reasons to map it directly or estimate its length by terminal redundancy measurements. However, with another class of stable diploids isolated in this laboratory it has been possible to locate precisely the associated stabilizing deletion (see succeeding paper).

A quite different method of conferring stability has been proposed by Adelberg & Bergquist (1972) to explain the properties of an *E. coli* strain in which an F-*gal* plasmid has been stably integrated into the host chromosome. In this scheme one arm of the duplication is inverted. This ensures stability but has the additional consequence that the duplication cannot be converted into a segregating structure by any type of genetic interaction, so the model is not applicable to the situation arising with the stable rII diploids. Stabilization of quite a different nature can be brought about by introducing duplications into a *rec⁻* genetic background. An interesting example of this type has recently been reported by N. Glansdorff (personal communication) for duplications covering the arginine region of *E. coli*. These duplications originally are unstable, but when transferred into bacterial hosts containing the *rec A* mutation they no longer generate any segregants. Although these different types of experiment demonstrate that stability of duplications in micro-organisms can be successfully brought about both by structural and genetic methods it is doubtful whether any of these schemes are directly applicable to the situation in diploid organisms in which chromosomes often appear to contain extended regions of repeated DNA sequences. Perhaps some quite different mechanism of ensuring stability applies in these cases which will only be appreciated when the basic processes involved in genetic recombination are more fully understood.

The nine new deletions we identified have not yet been mapped, although some are known to protrude into the duplication and so must be located near the rII region. Electron microscope heteroduplex studies similar to those of Bujard, Mazaitis & Bautz (1970) are now in progress. The three deletions whose lengths were estimated by terminal redundancy measurements are all longer than 2000 base pairs. It is easy to see that the selective procedure isolates deletions with a particular minimum length. The procedure starts with a particular s1231/3157 diploid, D, which has a high segregation frequency and thus possesses a long duplication. The replacement of 3157 by the small deletion 196 in this diploid further increases the length of the duplication and removes the terminal redundancy of the resulting phage which are then non-viable. New compensating deletions are then selected which restored viability and must, therefore, have a definite minimum length. As the replacement of 3157 by 638 in D results in non-viable phage the minimum length of the new deletions must be greater than the difference in the lengths of 638 and 196, which is about 1000 base pairs. If smaller deletions are required the selective process can easily be modified by isolating those compensating deletions which restore viability to the s1231/638 diploids derived from D.

The actual origin of the stabilizing deletion *sd*1 is of interest. The strain *hd*4 was originally isolated from the s1231/3157 diploid D as a presumed s1231/s1231 homozygous diploid. In some way, then, *hd*4 lost its homozygosity and acquired

*sd*1 in the second arm of the duplication. It is relevant in this connexion that s1231/638 diploids derived from D are non-viable. As s1231 and 638 are roughly of equal length it would be expected that s1231/s1231 diploids derived from D are also non-viable. Because *sd*1 is longer than s1231 this might explain why the s1231/*sd*1 diploids are viable and may therefore be found.
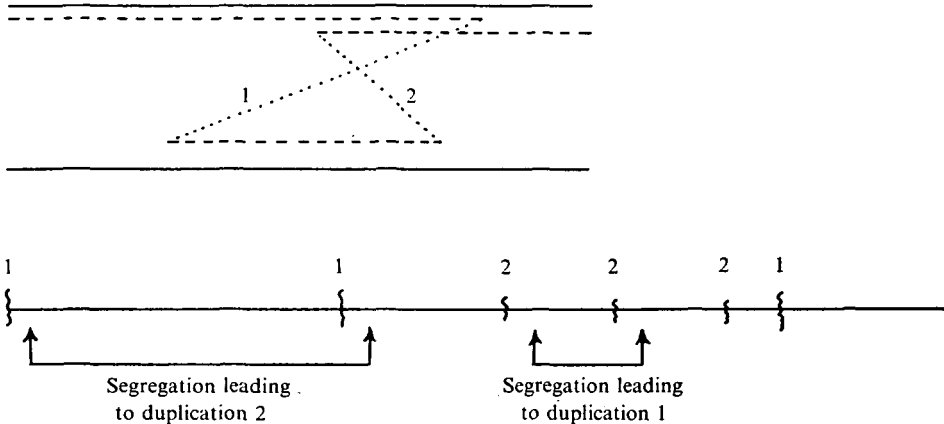

Fig. 7. Formation and structure of a *dudu*. Illegitimate recombination events 1 and 2 (dotted lines) lead to a structure in which one arm of duplication 1 includes the duplication 2.

We have found evidence for a number of novel structures involving repeated gene sequences. One of these is the *dudu;* or 'duplication of a duplication'. This was postulated in the first instance because both haploid and diploid segregants were isolated from certain recombinant phage. We would like to consider the properties of these *dudus* in relation to another as yet unexplained observation noted with the rII diploids. If one isolates spontaneous diploids the selection plaques often contain more than one type of diploid (Weil *et al.* 1965). In view of the rare occurrence of diploids this result is unexpected. A possible explanation of this finding is that the different duplications in the selection plaque arise from one initial *dudu*. Fig. 7 shows how this could occur. Segregation within the *dudu* would then lead to the production of two different duplications. The formation of the *dudu* requires two illegitimate recombination events at nearby places, which means that localized high negative interference for these events must be assumed. This proposal is different from that put forward by Symonds *et al.* (1972), who assume that the formation of the diploids is preceded by a rare phage mutation which alters a recombination enzyme in such a way that frequent mispairing occurs between phage genomes. This would also lead to high negative interference between illegitimate recombination events, but this would not be localized. Both models can explain the observed correlation between the formation of different duplications, and also the correlation between the formation of duplications and deletions observed by Weil & Terzaghi (1970). These authors suppose that the latter correlation rests on the frequent need for a new deletion to restore the ter-

minal redundancy removed by the insertion of the duplication. Such a model cannot explain the former correlation however. Moreover, it cannot explain related phenomena observed in systems other than T4 which are not subject to limitations caused by terminal redundancy. For example, the episome F13 can integrate into the chromosome of *E. coli* W3747 although there are no known regions of homology between the episome and the chromosome. Episome integration under these circumstances often creates deletions near the sites of integration (Fan, 1969).

REFERENCES

ADELBERG, E. A. & BERGQUIST, P. (1972). The stabilization of episomal integration by genetic inversion: a general hypothesis. *Proceedings of the National Academy of Sciences, U.S.A.* **69**, 2061–2065.

BENZER, S. (1962). On the topography of the genetic fine structure. *Proceedings of the National Academy of Sciences, U.S.A.* **47**, 403–415.

BUJARD, H., MAZAITIS, A. J. & BAUTZ, E. K. F. (1970). The size of the rII region of bacteriophage T4. *Virology* **42**, 717–723.

DOVE, W. F. (1968). The extent of rII deletions in phage T4. *Genetical Research* **11**, 215–219.

DRAKE, J. W. (1963). Mutational activation of a cistron fragment. *Genetics* **48**, 767–773.

EDGAR, R. S. & WOOD, W. B. (1966). Morphogenesis of bacteriophage T4 in extracts of mutant-infected cells. *Proceedings of the National Academy of Sciences, U.S.A.* **55**, 498–505.

FAN, D. P. (1969). Deletions in limited homology recombination in *Escherichia coli. Genetics* **61**, 351–361.

OHNO, S. (1970). *Evolution by Gene Duplication*. Berlin, Heidelberg, New York: Springer-Verlag.

PARMA, D. H., INGRAHAM, L. J. & SNYDER, M. (1972). Tandem duplications of the rII region of bacteriophage T4D. *Genetics* **74**, 319–335.

SEDEROFF, R., BOLLE, A. & EPSTEIN, R. H. (1971). A method for the detection of specific T4 messenger RNAs by hybridization competition. *Virology* **45**, 440–455.

SYMONDS, N., VAN DEN ENDE, P., DURSTON, A. & WHITE, P. (1972). The structure of rII diploids of phage T4. *Molecular and General Genetics* **116**, 223–238.

TESSMAN, I. & HALL, D. H. (1967). Linkage of T4 genes controlling a series of steps in pyrimidine biosynthesis. *Virology* **31**, 442–448.

VAN DEN ENDE, P. & SYMONDS, N. (1972). The isolation and characterization of a T4 mutant partially defective in recombination. *Molecular and General Genetics* **116**, 239–247.

VAN DE VATE, C. (1972). A stable duplication as an intermediate in selection for deletion mutants of phage T4. *Report to the Royal Society*.

VAN DE VATE, C., VAN DEN ENDE, P. & SYMONDS, N. (1974). A stable duplication of the rII region of bacteriophage T4. *Genetical Research* **23**, 107–113.

WARNER, H. R., SNUSTAD, D. P., JORGENSEN, S. E. & KOERNER, J. F. (1970). Isolation of bacteriophage T4 mutants defective in the ability to degrade host deoxyribonucleic acid. *Journal of Virology* **5**, 700–708.

WARNER, H. R., SNUSTAD, D. P., KOERNER, J. F. & CHILDS, J. D. (1972). Identification and genetic characterization of mutants of bacteriophage T4 defective in the ability to induce exonuclease. *Virology* **9**, 399–407.

WEIL, J. & TERZAGHI, B. (1970). The correlated occurrence of duplications and deletions in phage T4. *Virology* **42**, 234–237.

WEIL, J., TERZAGHI, B. & CRASEMANN, H. (1965). Partial diploidy in phage T4. *Genetics* **52**, 683–693.