JOURNAL OF BACTERIOLOGY, Oct. 1969, p. 269-275
Vol. 100, No. 1

Evolution of a Site of Specific Genetic Homology on the Chromosome of *Escherichia coli*

PHYLLIS L. KAHN

College of Biological Sciences, Department of Genetics and Cell Biology, University of Minnesota, St. Paul, Minnesota 55101

Received for publication 7 July 1969

Integration of the factors F_v and F into the chromosome of a substrain of *Escherichia coli* K-12 has been studied. The F_v factor is a fertility factor derived from Col V, lacking the ability to govern the production of colicin V. The derivatives of an Hfr_v (Hfr isolated from a V colicinogenic parent) strain, PK2 (initially isolated from C600 V^+), were shown to retain a unique bidirectional sex factor affinity locus between *recA* and *pheA*. This site shows no affinity for the *E. coli* K-12 F factor as shown by inability to isolate Hfr strains with origins in this region from a parental strain containing a cytoplasmic F factor. However this area exhibits two regions of homology to the V colicinogenic factor. One gives rise to Hfr_v strains identical to the original Hfr_v strain, PK2, with an origin and polarity of transfer designated *pheA*-CC injecting markers in the order *pheA-his-trp-pro*. The second gives rise to strains apparently originating at the same site but with reverse polarity designated *recA*-C, transferring markers in the order *recA-thyA-str-xyl*. For strains possessing the F_v factor only the second homology is apparent. A model for the evolution of these strains is presented.

The isolation of Hfr_v strains from a V^+ parent (Col V-K94) has been demonstrated previously (10). All the Hfr_v strains isolated from C600 (Col V), χ402 (Col V), or W1485 (Col V) showed the same origin and polarity of transfer; these strains were referred to as *xyl*-C (injecting markers in the clockwise order *xyl-pro-his-str*) with the exception of one, PK2. This strain was shown to have a counterclockwise (CC) direction of injection and an origin between *thyA* and *aroC*. It was also shown that the unique integration or Col V homology site of this strain remained a property of its chromosome, rather than the episome, when the Col V factor reverted to the cytoplasmic state (PK201), and was subsequently removed (PK211). The derivation of strain PK202 was also described previously (10). In this strain, derived from PK2, the genes for colicin V production are integrated into the bacterial chromosome and the sexuality properties are lost.

The present paper presents a more precise mapping of the origin of PK2 (between *recA* and *pheA*, designated *pheA*-CC) and the isolation of a new set of Hfr_v strains derived from PK211 or PK202, with their origins at the same locus but injecting the chromosome with inverted polarity, designated *recA*-C. A possible model for the chromosome of PK2 and a discussion of the unique homologies of the chromosomes of these strains with various episomes will be presented.

MATERIALS AND METHODS

The bacterial strains and their derivatives used in this study are described in Tables 1 and 2 and in an earlier report (10). The culture conditions and methods such as *N*-methyl-*N'*-nitro-*N*-nitrosoguanidine (NTG) mutagenesis, the isolation of Hfr strains and the curing of cytoplasmic factors from bacterial strains have been described previously (10, 11). Inversion Hfr strains were isolated from PK2 by picking clones showing either high *thy* recombination or low *his* recombination. The genotype symbols are those customarily accepted (4, 13).

RESULTS

Isolation and mapping of origins and directions of Hfr strains. The spectrum of origins of Hfr or Hfr_v strains isolated was used as an indication of genetic homology between chromosome and episome. Initially, the formation of Hfr strains from PK202 (10) was studied. This strain was isolated after the addition of an F-*lac* factor to PK2, presumably as a result of a fertility factor immunity conflict (11). The F-*lac* factor was subsequently lost and PK202 produces colicin V, does not ferment lactose, is resistant to phage MS2 and is unable to recombine or transfer the

TABLE 1. *List of strains*

Strain	Composition or derivation
AT2092	*phe A-pur Fl-his⁻-arg H-str*ʳ-T1ʳ-T6ʳ-*lac⁻-mal⁻-xyl⁻-mtl⁻*-Vʳ [a]
AB2495	*thr⁻-leu⁻-thi⁻-trp⁻-pro⁻-his⁻-arg⁻-thy⁻-lac⁻-gal⁻-ara⁻-xyl⁻-mtl⁻-str*ʳ-T6ʳVʳ
AB2463	*thr⁻-leu⁻-thi⁻-pro⁻-his⁻-arg⁻-lac⁻-gal⁻-ara⁻-xyl⁻-mtl⁻-str*ʳ-T6ʳ-Elʳ-Vʳ
PK2465	Same as AB2463, but with *thy⁻*. Derived from AB2463 by trimethoprim mutagenesis (5).
C600 [b]	*thr⁻-leu⁻-thi⁻-lac⁻*-Vʳ-Elʳ-*Azʳ-str*ʳ
C600 F^+	
C600 F_v^+	(10)
C600 (Col V)	(11)
PK2	Hfr_v (10)
PK201	V^+ Hfr^- segregant of PK2 (10)
PK202	(10)
PK205	PK202 F^+ (10)
PK206	PK202 F_v^+
PK211	Acridine orange curing of PK201 (10)
PK211 F^+	
PK211 (Col V)	
PK211 F_v^+	
PKm128	Strain with a mutant Col V factor characterized by a larger than normal colicin V inhibition zone
PK207	PK202 (Col V-PKm128)
C600 (Col V-PK2061)	This paper

[a] Received from K. B. Low.
[b] Following strains are all derived from C600.

Col V factor (10). As listed in Table 1, PK205 arises from the addition of an F factor to PK202; PK206 from the addition of an F_v factor (a fertility factor initially derived from the Col V factor); and PK207 from the addition of a mutated V colicinogenic factor, V-PKm128 (distinguishable because of its larger colicin V inhibition zone). PK211 was obtained from acridine orange curing of PK201 (10) to obtain an F^- strain having the chromosome of PK2 without any cytoplasmic particle.

The Hfr strains isolated are shown in Fig. 1. Table 2 lists their origins, directions of injection and parental derivation. Strains with origin identical to one indicated have been omitted from Fig. 1. In addition, strains apparently identical to ones listed in Table 2 have often not been cited. The use of two additional markers *pheA* (in the recipient AT2092, received from K. B. Low) and *recA* in PK2465 (Table 1) permitted a more precise mapping of the origin of PK2 and other Hfr strains, originally only localized between *thyA* and *aroC*. Table 3 gives the data for the mapping of the origins with respect to the *recA* and *pheA* genes. Clark (3) has shown that an Hfr strain with a low recombination deficiency index (defined as the ratio of the number of recombinants formed from a Rec^+ recipient to the number formed with a Rec^- recipient for a given marker from a given donor) injects Rec^+ into a *recA* strain as an early marker while strains with a high recombination deficiency index do not. Mapping of the origins and directions of the Hfr strains with respect to *thy* and *aroC* gave the information that strains, such as PK37, inject *thy* as an early marker and must, therefore, inject Rec^+ early. Thus, a recombination deficiency as high as 66.7 (from PK37) should be considered low.

As will be discussed at greater length later, the data indicate no effect of the PK2 chromosome on the integration of the F factor. The sites of integration of the F factor in the Hfr strains formed from PK205 seem completely unaffected by the presence of the partial Col V factor, as can be seen by comparing Fig. 1 with the pattern of F factor integration obtained earlier (10). Concerted efforts to obtain an F factor integrated in the region between *pheA* and *recA* were not successful although Hfr strains with nearby insertion points were isolated from both PK205 and C600 F^+ (Table 2 and Fig. 1). In addition, preliminary results indicate that PK26, the only F^+ strain previously reported (10) as integrating at the same site as most Hfr_v strains (between *mal* and *xyl*), has been mismapped and actually integrates between *thy* and *str*. Therefore, although the F factor and the Col V factor have much in common (11), they appear to differ strongly in their integration specificities.

TABLE 2. *List of Hfr strains*

Strain	Parent	Origin	Direction
PK2	C600 (Col V)	*pheA*[a]	CC[b]
PK16	C600 (Col V)	*recA*	CC
PK38	C600 F^+	*aroC*	CC
PK36	C600 F^+	*recA*	CC
PK37	C600 F^+	*thyA*	CC
PK2061	PK206 = PK202 F_v^+	*recA*	C
PK2062	PK206	*xyl*	C
PK2063	PK206	*recA*	C
PK2064	PK206	*xyl*	C
PK2065	PK206	*xyl*	C
PK2066	PK206	*recA*	C
PK2051	PK205 = PK202 F^+	*his*	CC
PK2052	PK205	*pro*	CC
PK2053	PK205	*recA*	CC
PK2054	PK205	*arg*	CC
PK2055	PK205	*his*	CC
PK2056	PK205	*arg*	CC
PK2057	PK205	*his*	CC
PK2058	PK205	*ara*	C
PK2059	PK205	*pro*	CC
PK20510	PK205	*pheA*	C
PK2071	PK207 = PK202 (Col V-128)	*xyl*	C
PK2072	PK207	*recA*	C
PK2073	PK207	*pheA*	CC
PK2081	PK2	*recA*	C
PK2111	PK211 F_v^+	*xyl*	C
PK2112	PK211 F_v^+	*recA*	C
PK2113	PK211 (Col V)	*recA*	C
PK2114	PK211 (Col V)	*pheA*	CC
PK2011	PK201	*pheA*	CC
PK2012	PK201	*pheA*	CC

[a] Proximal marker injected of those tested.
[b] C, clockwise; CC, counterclockwise as shown in this paper, Fig. 1 and (10, 12).

Quite different results were obtained upon the infection of PK202 with other episomes. From PK206 (PK202 F_v^+), two types of Hfr strains were isolated (Fig. 1). The PK2062 type has the typical location found from V^+ or F_v^+ strains; however, the other, found with approximately equal frequency, has an origin in the same region as PK2, but transfers the chromosome with the opposite polarity (illustrated by PK2061). However, from PK207 (PK202 with the V colicinogenic factor from PKm128), three possible integration sites were demonstrated: the *xyl*-C site (as PK2071), the *recA*-C site (PK2072), and the *pheA*-CC site (PK2073) of PK2.

Hfr strains were also isolated from PK201 and from PK211 with various episomes. It has been shown previously (10) that the specificity for an integration site leading to *aroC*-CC (now *pheA*-CC) Hfr strains and high *his* recombination from a donor with a cytoplasmic Col V factor

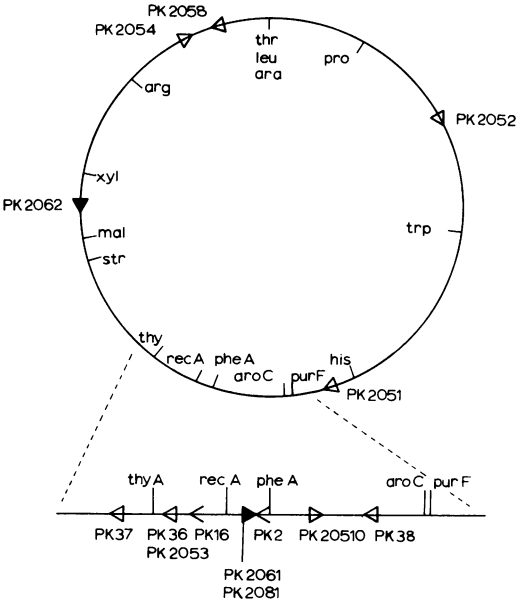

FIG. 1. *Linkage map showing the point of origin of chromosome transfer for Hfr and Hfr$_v$ strains isolated from E. coli K-12.*

Symbols: ▲, parental F^+ strain; <, parental Col V strain; △, parental F_v^+ strain.

TABLE 3. *Mapping of the origins of Hfr strains with respect to recA and pheA*

Donor	No. of recombinants[a]			Recombination deficiency index[b]
	pheA	Rec^+ recip.	Rec^- recip.	
PK2	47.5	127	0.008	1.59×10^4
PK2011	344	82	0.005	1.64×10^4
PK2012	156	47	0.008	5.9×10^3
PK2053	72.9	130	13.0	10
PK38	0.072	42	0.018	2.3×10^3
PK36	82.9	50	2.34	21.4
PK37		82	1.23	66.7
PK2061	0.24	196	16.2	12.1
PK2063	0.12	154	30.5	5.05
PK2064		121	0.0051	2.4×10^4
PK2066	0.105	201	13.5	14.9
PK2112	0.071	192	20.2	9.5
PK2081	0.153	333	5.0	66.5

[a] A number of 100 is approximately equivalent to 1×10^7 recombinants/ml or .4 recombinants/initial donor cell. The Rec^+ recipient was AB2495. The Rec^- recipient was PK2465. The marker used was *his*, *thy*, or *arg* whichever was closest to the origin for each donor strain.
[b] Recombination deficiency index = Rec^+/Rec^- recombinants for a given marker (3).

resided with the chromosome rather than with the episome of PK201. From the isolation of Hfr strains, it can be seen that PK211 or PK201 has two major regions of homology with the Col V factor, one giving *recA*-C type Hfr strains and the other giving *pheA*-CC type Hfr strains. However, in the case of PK211 F_v^+, only the site leading to *recA*-C integration is apparent in this area, along with the usual *xyl*-C F_v or Col V specific site. In addition, it should be noted that inversion Hfr strains (i.e., PK2081) can be isolated from PK2 itself at a fairly high frequency (0.5%), whenever selection for this type of Hfr is made.

Location of the ability to produce colicin V in PK205. Earlier (10) it was shown that the genes for colicin V production in PK205 have a chromosomal location between *thy* and *aroC*. With the use of the recipient AT2092, close linkage to the *pheA* gene was shown (Table 4). This was expected as the origin of PK2 was shown to be within 1 1/2 min of *pheA*.

Fertility of F^+, V^+, and F_v^+ strains. It has been shown (10) that the recombination rate for various markers from donors containing episomes and the type of Hfr most often isolated usually exhibit a similar bias. This gives a preliminary check on the affinity of an episome for a particular chromosomal site easier to obtain, compared to data obtained from the isolation of Hfr strains. Table 5 shows recombination data of donor strains derived from C600 or PK2, containing different fertility factors. All the strains derived from PK2 show an increase in the number of *his* and *thy* recombinants, compared to C600 (Col V); the most striking effects are on the *his* recombinants from PK207 and, to a lesser extent, the number of *thy* recombinants from PK206. It should also be noted that the Col V factor from PK2 or PK2061 shows chromosomal affinity identical to the ordinary Col V factor, once separated from the chromosome of a PK2-derived strain. The conclusions on genetic homologies drawn from the recombination data of strains containing the different episomes are seen to agree with those reached from the study of episome integration sites

DISCUSSION

The data presented indicate the existence of some very specific genetic homologies between chromosome and episome in the clones derived from PK2.

TABLE 4. *Transfer of colicinogeny to recombinants selected for listed markers*[a]

Recombinant for	No. of recombinants		
	Total tested	Colicinogenic	
		No.[b]	%
pheA	90	86	96
purF	90	47	52
his	90	12	13
arg	90	0	0

[a] Donor PK205, recipient AT2092.

[b] None of the V colicinogenic strains tested (4 *his* recombinants, 10 *phe* recombinants and 10 *pur* recombinants) were able to transfer the Col V factor at high frequency.

TABLE 5. *Recombination of strains derived from C600 containing various episomes*

Donor	Episome	*arg*		*pro*		*his*		*thy*	
		No.[a]	Percentage of *arg*	No.	Percentage of *arg*	No.	Percentage of *arg*	No.	Percentage of *arg*
C600	F^+	115	100	92	80	136	118	30.7	27
C600	V^+	101	100	45	45	12.3	12	4.9	4.9
C600	F_v^+	184	100	66	36	14.1	7.7	3.7	2
C600	(Col V-PK2)	110	100	68.6	62	12	11	3.8	3.5
PK201	(Col V-PK2)	114	100	75	66	75	66	71	62
PK211	(Col V)	110	100	33.9	31	37.9	35	36.9	34
PK211	F^+	93	100	95	102	128	137	37	40
PK211	F_v^+	329	100	134	41	14.2	4.3	79	24
PK205	F^+	230	100	213	93	337	146	130	57
PK206	F_v^+	296	100	150	51	190	64	650	219
PK207	(Col V-PKm128)	201	100	148	74	4960	2470	377	187
C600	(Col V-PK2061)	271	100	85	31	21.3	7.9	16.6	6.4

[a] A number of 100 would be equivalent to a recombination rate of approximately 10^{-4} per initial donor cell.

Two strains derived in distinct ways from PK2 were used in this study. PK202, having only the V-producing genes and none or little of the fertility determinants of the V colicinogenic factor, was isolated presumably as the result of a fertility factor-immunity conflict (11). When an F factor is added to this cell (PK205) and the episome-chromosome interaction is studied either through Hfr isolation (Table 2 and Figure 1) or recombination for different markers (Table 5), no unique genetic homology between chromosome and episome can be demonstrated. The spectrum of the origins of Hfr strains isolated is very similar to that found earlier for several *E. coli* K-12 F^+ donors (10).

However, when either Col V-PKm128 or the F_v factor is used as the episome in PK207 and PK206, respectively, an entirely different pattern results. Three types of Hfr strains are isolated from PK207: *xyl*-C as PK2071, *recA*-C as PK2072, and *pheA*-CC as PK2073. The recombination data of Table 5 show that the strongest region of genetic homology or the most frequent chromosome-episome interaction occurs in the last of the three. However, in PK206, this site of Hfr origin is lacking; only the *xyl*-C origin (as in PK2062) or the *recA*-C origin (as in PK2061) is evident.

Similar results were obtained with the other basic strain used, PK211. This strain was obtained from acridine orange curing of the Col V factor from PK201, a V^+ Hfr^- revertant of PK2. Thus, it has the chromosome of PK2, without any of the properties governed by Col V. Again, no specific episome-chromosome interaction is observed with PK211 F^+ for either Hfr formation or recombination. PK211 (Col V) shows the three chromosomal regions of genetic homology with Col V typical of PK207 with one difference. The very strong homology leading to the *pheA*-CC Hfr strain is now reduced and is of the same order as the other two. This is to be expected as PK207 has the deoxyribonucleic acid base sequence to produce colicin V on both episome and chromosome, and PK211 (Col V) has it only on the episome. The episomes of PK211 F_v^+ and PK206 completely lack this region of homology.

A model incorporating these results is presented in Fig. 2. It is to be emphasized that this model is to demonstrate only a possible evolutionary scheme for the observed site homologies.

The first step postulates a tandem duplication of Col V in the chromosome of PK2 between *recA* and *pheA*. This was never isolated, but such duplications have been observed elsewhere in the *E. coli* chromosome (8, 9). In addition, concatenated or multiple forms of a colicinogenic factor (Col E_1) have also been observed (7). In steps 2–2a, an inversion of the F_v portion of one of the duplicated genomes is postulated to yield step 3, a schematic representation of the chromosome of PK2. As shown in step 2, it is assumed that this inversion inactivated parts of both the fertility determinant segment (*def*) and the colicin V production determinant (*abc*), respectively *f* and *c*. There is no known a priori reason for this step to occur, but it is a convenient way to arrive at step 3.

From step 3 on, the experimental evidence and theoretical reasoning are on somewhat firmer ground. The chromosome of PK2 must contain a fully functional integrated V colicinogenic factor (*abcdef*), plus two additional regions of homology to Col V (*ab* and *ed*) and one region of homology to the F_v factor (*ed*). The segment *ed* is inverted with respect to its normal relationship to *ab*.

From this postulated structure of PK2, the discussion of the effect of nucleotide sequence redundancies presented by Thomas (14) can be utilized. Pairing of the intact Col V factor with the *ab* segment (step 4) leads to the creation of episomes as in strain PK201 (step 4a). Subsequent removal of Col V (step 4b) gives strain PK211 which still retains the normal and the inverted regions of homology, but no longer is fertile or produces colicin V. This has been called homodirectional pairing by Berg and Curtiss (1). Pairing of the intact Col V factor with the inverted segment (steps 5–5a) leads to a completely inverted Col V factor (step 5b) in the chromosome (14). This has been called heterodirectional pairing (1).

A partial inversion (step 6–6a) similar to that shown in step 2–2a (again without any supportive evidence) results in the formation of PK202 (step 6b) which produces colicin V (implying complete activity of *abc*), has a very strong forward homology to Col V (as shown in Table 5 by the recombination of PK207) and an inverted homology to F_v (shown in Table 5 by the recombining ability of PK206).

It should also be pointed out that the ability to produce colicin V is never found to enter as an early marker in any of the Hfr strains derived from PK2, including the inversion strains. Therefore, to maintain proper F_v and colicin V function, an inversion of the entire Col V factor seems to be necessary. Thomas (14) points out that inversions within a gene will probably be nonfunctional. For an inversion to be functional, it must contain a full operon along with the site for initiation of transcription. The partial in-

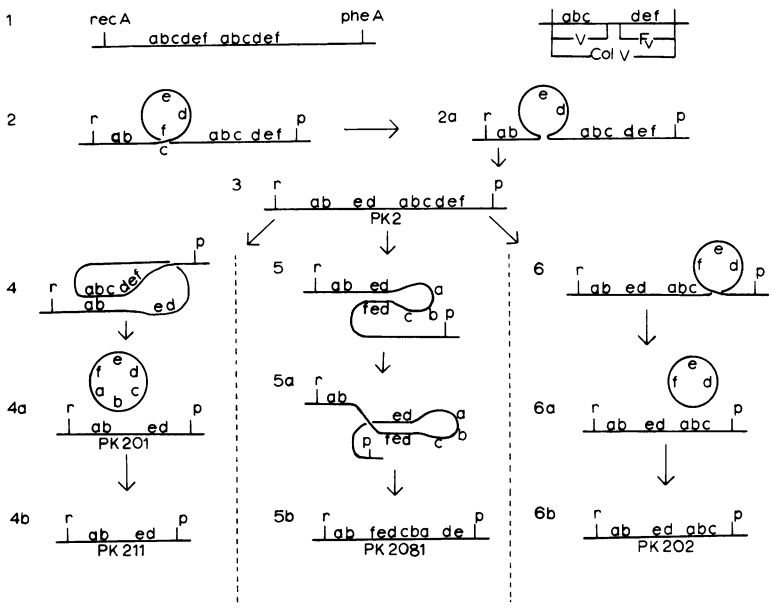

FIG. 2. *Symbols: abcdef is a schematic representation of Col V, abc representing the genetic determinant of colicin V production and def the genetic determinant of Col V-associated fertility (F_v); the length of these segments does not represent any known distance and three letters are used for each solely to indicate directionality; r represents the recA locus at 51.5 min and p represents the pheA locus at 50 min (13). Steps: 1. Postulated first step, consisting of tandem duplication of Col V in the E. coli chromosome between recA and pheA. 2–2a. Postulated inversion of the F_v portion of one of the duplicated genomes to yield 3. 3. Postulated structure of the chromosome of PK2 which contains one functional Col V factor and two inactive parts of the second Col V factor, one of which is inverted. 4–4a. Pairing with the repeated noninverted segment (14) leading to episome formation in 4a (PK201). 4b. Loss of the episome leaving a chromosome with a normal and inverted region of homology to Col V and an inverted homology to F_v (PK211). 5–5a. Pairing with the repeated inverted segment (14) to give a completely inverted Col V in the chromosome (PK2081). 6–6a. A step similar to 2, resulting in the formation of PK202 (6b) which produces colicin V (implying complete activity of abc), has a very strong forward homology to Col V and an inverted homology to F_v and, also, Col V.*

versions postulated in steps 2 and 6 of Fig. 2 both lead to loss of some function.

As stated previously (10), the state of the Col V factor in the Hfr_v strains is not completely understood. The transfer of the Col V factor associated with distal markers in these experiments seems lower than the reports of others (6). It is possible that the Hfr_v strains represent an intermediate state of association between episome and chromosome. For simplicity, the model of Campbell (2) for insertion of episomes into the chromosome has been assumed. There are clearly further experiments to be done with this system. For example, the homology relationships imply a deletion of material occurring in the transition from Col V to F_v. Experiments to determine the size of these factors are now in progress. The model presented implies that the chromosomal order of *recA*-Col V-F_v-*pheA* in PK2 should be *recA*-F_v-Col V-*pheA* in PK2061 or PK2081. This can be tested using P1 transduction, since the distances involved should be within the range of distances studied by marker cotransduction (13).

ACKNOWLEDGMENTS

Support was provided by Public Health Service grant GM14316 from the division of General Medicine to Leon A. Snyder. I wish to thank Sally Berg for technical assistance.

LITERATURE CITED

1. Berg, C., and R. Curtiss. 1967. Transposition derivatives of an Hfr strain of *E. coli* K-12. Genetics **56**:503–525.
2. Campbell, A. M. 1962. Episomes. Advan. Genet. **11**:101–145.
3. Clark, A. J. 1967. The beginning of a genetic analysis of recombination proficiency. J. Cell. Physiol. **70**(Suppl. 1):165–180.
4. Demerec, M., E. A. Adelberg, A. J. Clark, and P. E. Hartman. 1966. A proposal for a uniform nomenclature in bacterial genetics. Genetics **54**:61–76.
5. Dubnau, E., and W. K. Maas. 1968. Inhibition of replication of an F′ *lac* episome in Hfr cells of *Escherichia coli*. J. Bacteriol. **95**:531–539.
6. Fredericq, P. 1965. Genetics of colicinogenic factors. Zentrabl. Bakteriol. Parasitenk. Abt. Orig. **196**:142–151.
7. Goebel, W., and D. R. Helinski. 1968. Generation of higher

multiple circular DNA forms in bacteria. Proc. Nat. Acad. Sci. U. S. A. 61:1406–1413.
8. Hill, C. W., J. Foulds, L. Soll, and P. Berg. 1969. Instability of a missense suppressor resulting from a duplication of genetic material. J. Mol. Biol. 39:563–581.
9. Horiuchi, T., S. Horiuchi, and A. Novick. 1963. The genetic basis of hyper-synthesis of β-galactosidase. Genetics 48:157–169.
10. Kahn, P. L. 1968. Isolation of high-frequency recombining strains from *Escherichia coli* containing the V colicinogenic factor. J. Bacteriol. 96:205–214.
11. Kahn, P. L., and D. R. Helinski. 1964. Relationship between colicinogenic factors E_1 and V and an F factor in *Escherichia coli*. J. Bacteriol. 88:1573–1579.
12. Matney, T. S., E. P. Goldschmidt, N. S. Erwin, and R. A. Scroggs. 1964. A preliminary map of genomic sites for F attachment in *Escherichia coli* K-12. Biochem. Biophys. Res. Commun. 17:278–281.
13. Taylor, A. L., and C. D. Trotter. 1967. Revised linkage map of *Escherichia coli*. Bacteriol. Rev. 31:332–353.
14. Thomas, C. A. 1966. Recombination of DNA molecules, p. 315–337. *In* J. N. Davidson and W. E. Cohn (ed.), Progress in nucleic acid research and molecular biology, vol. 5. Academic Press Inc., New York.