Volume 17 Number 13 1989 Nucleic Acids Research

Cosmid-derived map of *E.coli* strain BHB2600 in comparison to the map of strain W3110

Rainer P.Birkenbihl and Walter Vielmetter

Institut für Genetik der Universität zu Köln, Weyertal 121, D-5000 Köln 41, FRG

Received April 7, 1989; Revised and Accepted June 12, 1989

ABSTRACT

A physical map for the genome of *E.coli* K12 strain BHB2600 was constructed by use of 570 cloned DNA elements (CDEs) withdrawn from a cosmid library. Dot blot hybridisation was applied to establish contig interrelations with subsequent fine mapping achieved by analysis of EcoR1 restriction patterns on Southern blots. The derived map covers nearly 95% of the *E.coli* genome resulting in 12 minor gaps. It may be compared to the almost complete map for strain W3110 of Kohara et al. (1). Except for one tiny gap (lpp,36.5') remaining gaps in BHB2600 do not coincide with those in W3110 so that both maps complement each other establishing an essentially complete clone represented map. Besides numerous minute differences (site and fragment gains and losses) both strains harbour at differing positions extended rearrangements flanked by mutually inverted repetitive elements, in our case insertion elements (IS1 and IS5).

INTRODUCTION

Numerous studies have attempted to establish physical maps of pro- and eucaryotic genomes. Progress has been reported for the successful construction of physical genomic maps such as the now close to complete map for the 4700kb long *E.coli* genome (1, 2, 3, 4) and partial maps for the 3 fold larger genome of the yeast *Saccharomyces* (5) and the 20fold larger genome of the nematode *Caenorhabditis* (6, 7). The aim of constructing such maps is to obtain quick physical access to any desired region along a given genome by selection of the desired clones from an ordered clone collection.

The general way to construct a physical genomic map is to search for overlaps between neighbored cloned DNA-elements, called contigs. Contigs grow and fuse together upon further random search or by walking procedures into larger contiguous and ordered map segments which might be aligned with genes positioned on an already existing classical map. One of the most preferred methods to detect overlaps between cloned elements is to compare in (computer aided) search programs combined band patterns of various restriction fragments (1, 2, 3, 5, 7). Only one recent attempt uses hybridisation as a basic mapping tool(4). In most other cases hybridisation techniques for overlap detection are used only at an advanced stage of mapping, mainly to fill gaps in the map (1, 6). In this way the so far most complete physical map of the K12 strain W 3110 of *E.coli* was constructed by use of 3400 Lambda-clones leaving 8 minor gaps beyond final completion (1). A set of 3500 cosmid clones in another approach provided material to close two of the remaining gaps (2).

We present here a close to complete map of the *E.coli* K12 strain BHB 2600. As a basic method for contig search we applied hybridisation by dot blot analysis while restriction patterns as a comparative tool were consulted for the later ordering of map elements. By

this 'inverted' strategy <600 cosmid clones were required to saturate the map up to 95%. With exception of one gap we were able to close the remaining gaps in the physical map constructed by Kohara et al.(1). Strain specific differences between the two maps will be discussed.

MATERIALS AND METHODS

Cosmid vector: pJB8 (47).

Cloned bacteria strain: E.coli K12 BHB2600 ($supE^+supF^+rk^-mk^+met^-$), derived from B.Hohn 1980.

Library construction was carried out according to Ish-Horowitz and Burke (47). BHB2600 DNA was isolated and partially digested with Sau3A. Four fractions, taken after 2, 4, 5 and 6 min from a single reaction in the size range between 35 and 50 kb, were dephosphorylated with alkaline phosphatase and ligated into the BamH1 site of cosmid vector pJB8. The in vitro packaged DNA was transduced into hoststrain HB101 and plated under conditions of ampicillin selection.

Preparation of cosmid clones: Picked colonies were grown to saturation in DYT medium containing 100 μg/ml ampilicin. Aliquots were stored as glycerol-stocks at −70°C. DNA was prepared using the alkaline lysis miniprep method (47).

Restriction digests were carried out according to the recommendations of the suppliers (Boehringer, Mannheim), EcoR1 digests with a twofold excess of enzyme.

Southern blots: Restricted DNA was separated on 0.6% agarose gels, 20×20×0.5 cm, 50 slots and blotted by alkaline transfer onto Gene Screen Plus membranes (48).

Dot blots: 200 ng of EcoR1 digested cosmid DNA, mixed with 1 μg sheared salmonsperm DNA was denaturated for 15 min in 150 μl 0.4M NaOH, 0.6M NaCl. The denatured DNA was distributed in dots onto Amersham Hybond N membranes by use of a micro-sample filtration manifold (Schleicher and Schüll). After drying the DNA was immobilized on the filters by UV irradiation (Herolab UVT2020) for five minutes.

Hybridisation and rehybridisation were performed according to Chomczynski and Qasba (48). The same procedure was applied for Southern blots and dot blots. DNAs were 32P-nicktranslated by use of a nicktranslation kit (Boehringer, Mannheim).

Vector competition: 1 μg vector DNA was mixed with 100 ng nicktranslated cosmid DNA in 100 μl 5×SSC, boiled for 10 min, quickchilled in icewater and incubated for 4 hours at 65°C before addition to the hybridisation mix.

CONSIDERATIONS FOR THE MAPPING PROCEDURE

Our aim here is to construct a fairly complete physical *E.coli* map employing the smallest feasible number of cloned *D*NA *e*lements (CDEs). The experimental scheme is outlined in Fig.1. For library construction *(1)* cosmids are chosen as vectors, known to insert CDE length close to 1% of the length of the *E.coli* genome so that about 500 randomly overlapping CDEs would suffice to saturate the map up to 99.5%. Propagation of cosmid inserted *E.coli*-DNA in *E.coli* is suspected to cause selective clone losses by gene dosage effects (2, 4). In order to obtain a reasonably wide CDE spectrum for a high map representation we tried to withdraw from the library cosmid containing bacterial colonies of all sizes including very slowly growing ones. Care was taken to start the mapping procedure with a 'clean' CDE working set *(2)*. After a first survey by restriction analysis *(3)* insert defective cosmids were rejected, DNA of weakly expressed minipreps replenished and host DNA contaminations recorded.

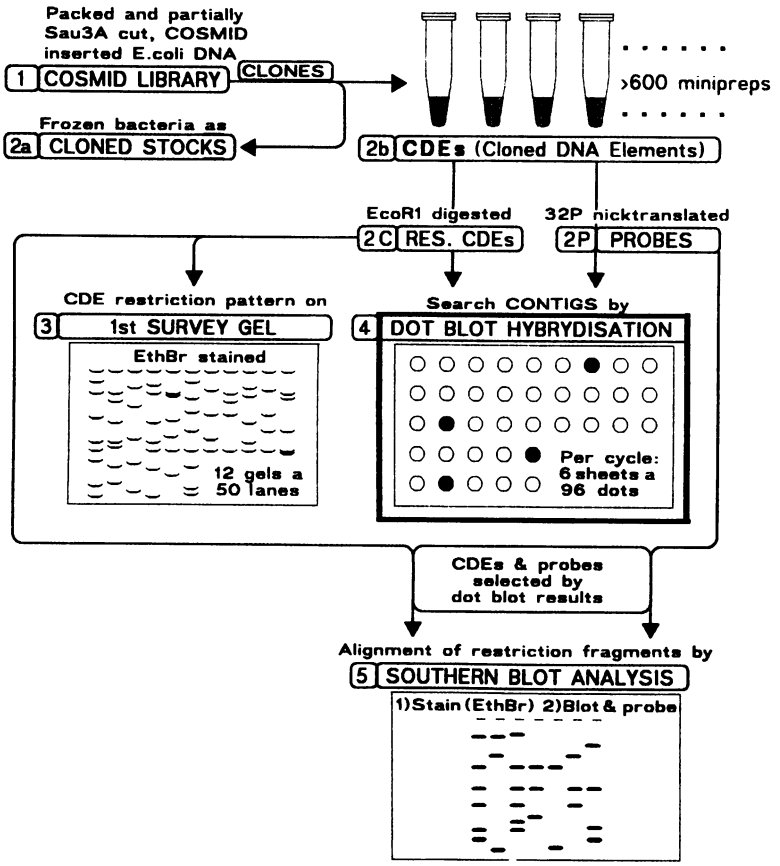


Fig.1. *Outline of the mapping procedure*: **1)** The *cosmid library* is constructed from DNA of *E.coli* K12/BHB 2600 partially Sau3A digested for various times, pooled and inserted into the BamHI site of cosmid vector PJB8 (47). **2)** >600 bacterial clones picked from the library supply the *cloned stock collection* **(2a)** and as minipreps the *set of cloned DNA elements (CDEs)* **(2b)** from which *EcoR1 digested CDEs* **(2C)** or *32P nick-translated probes* **(2P)** are derived. **3)** EthBr stained *1st survey Gels* (50 lanes) provide information on CDE restriction patterns to correct in minipreps anomalies concerning defective inserts, DNA concentration or host DNA contaminations. **4)** *Dot blot hybridisation* as the basis of mapping detects contigs of neighbored CDEs by overlap recognition. Overlap signals are generated by hybridisation of one labelled CDE probe per cycle against the whole (and at later cycles progressively reduced) CDE working set, distributed in dot patterns (96 dots) over several (30x) reusable nylon sheets. Vector signals are surpressed by a 100 fold excess of nonlabelled vector DNA. **5)** *Southern blot analysis* provides at a progressive stage of contig collection the information to align CDEs within higher order contigs and restriction fragments within CDEs. Complete restriction band patterns are compared by blot analysis with restriction fragments hybridizing with CDE probes of the same subset.

The main tool of our mapping procedure is dot blot hybridisation *(4)* which detects overlaps between 32P-nicktranslated CDE probes and members of the CDE working set *(2)*, distributed as a dot pattern on reusable nylon sheets. By use of a 'clean' DNA working set overlap signals can be recognized with high sensitivity and reliability. Per hybridisation cycle one overlapping CDE cluster (primary contig) of about 100kb average length can

Table 1. Cloned genetic markers used as probes for map correlation.
Map position refers to the classical linkage map (9). Ref.No.: indicates relevant publications and authors, who kindly provided probes.

No. in fig. 2	Gene	Map pos. [min]	Ref. No.	No. in fig. 2	Gene	Map pos. [min]	Ref. No.
1	dnaJ	0.3	10	18	glyA	54.9	27
2	folA	1.0	11	19	exbB	64.9	28
3	aroP	2.7	12	20	cysG	74.0	29
4	phoA	8.9	13	21	livH	75.9	30
5	adk	11.2	14	22	livB	82.2	31
6	gltA	16.5	15	23	sfrB	86.1	32
7	lit	25.4	16	24	lexA	91.8	33
8	rac	30.0	17	25	melA	93.4	34
9	kim	34.0	18	26	IS4	98.1	35
10	pheST	37.5	19				
11	motB	41.8	20	H	rrnH	5.2	36
12	flaA	42.7	21	G	rrnG	56.6	36
13	dcm	43.0	22	D	rrnD	72.2	36
14	hisG	44.0	23	A	rrnA	86.5	36
15	ompC	47.7	24	B	rrnB	89.8	36
16	dapA	53.1	25	E	rrnE	90.5	36
17	hisS	54.1	26				

be randomly drawn from the whole CDE working set. During the initial random search several cycles are run simultaneously. Guided by first survey gel patterns exclusively CDE probes from outside the existing contig pool are chosen in order to add only new, not yet identified contigs to the pool. With an increasing pool size probe signals common to differing first order contigs indicate fusion into second order contigs which in turn may proceed to higher order contigs.

Hybridisation derived contigs can be at best resolved to the CDE level. Exact alignment beyond this level is achieved by a one-enzyme (EcoR1) restriction analysis on Southern blots *(5)* probed with CDEs from the same contig subset. To align the singular array of restriction fragments CDEs are applied as a counter-array of independently interlocking elements. Disrupted restriction fragment residues at CDE termini are the decisive directional indicator, readily detectable by comparison with fragments of restricted whole *E.coli* DNA included in blot gels.

At a progressed stage of mapping systematically chosen hybridisation probes are applied: 1) At an extent of map growth which leaves merely smaller gaps in the map, dot blot probes are selected from CDEs in flanking positions at gap edges in order to bridge gaps by mini-walking. 2) A special set of probes from cloned material adjoined to gene positions on the classical *E.coli*-map (sources listed in Table 1) serves for comparative purposes and as an aid for the positioning of locally indetermined contig complexes.

RESULTS

The physical map of *E.coli* K12 strain BHB 2600 presented here (Fig.2) is constructed from 570 cosmid clones to approach 95% of the genome length. It is represented by a single restriction fragment array (8) with >600 EcoR1 fragments estimated for 100% map length. The comparably complete map deviced by Kohara et al. (1) for the K12 strain

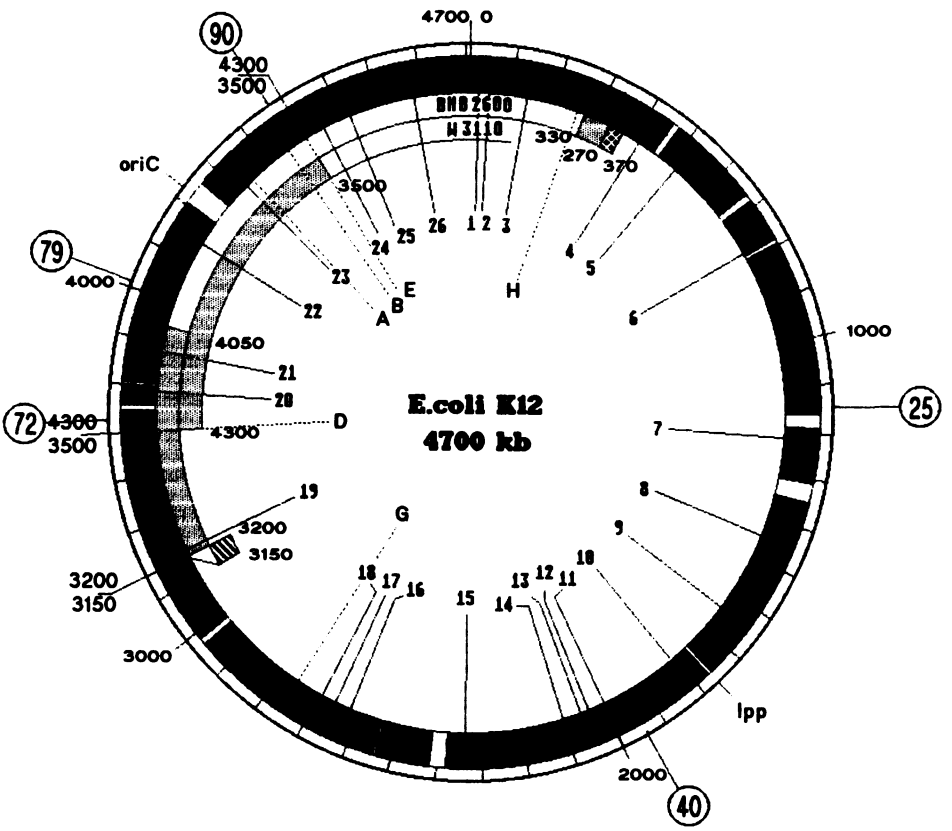
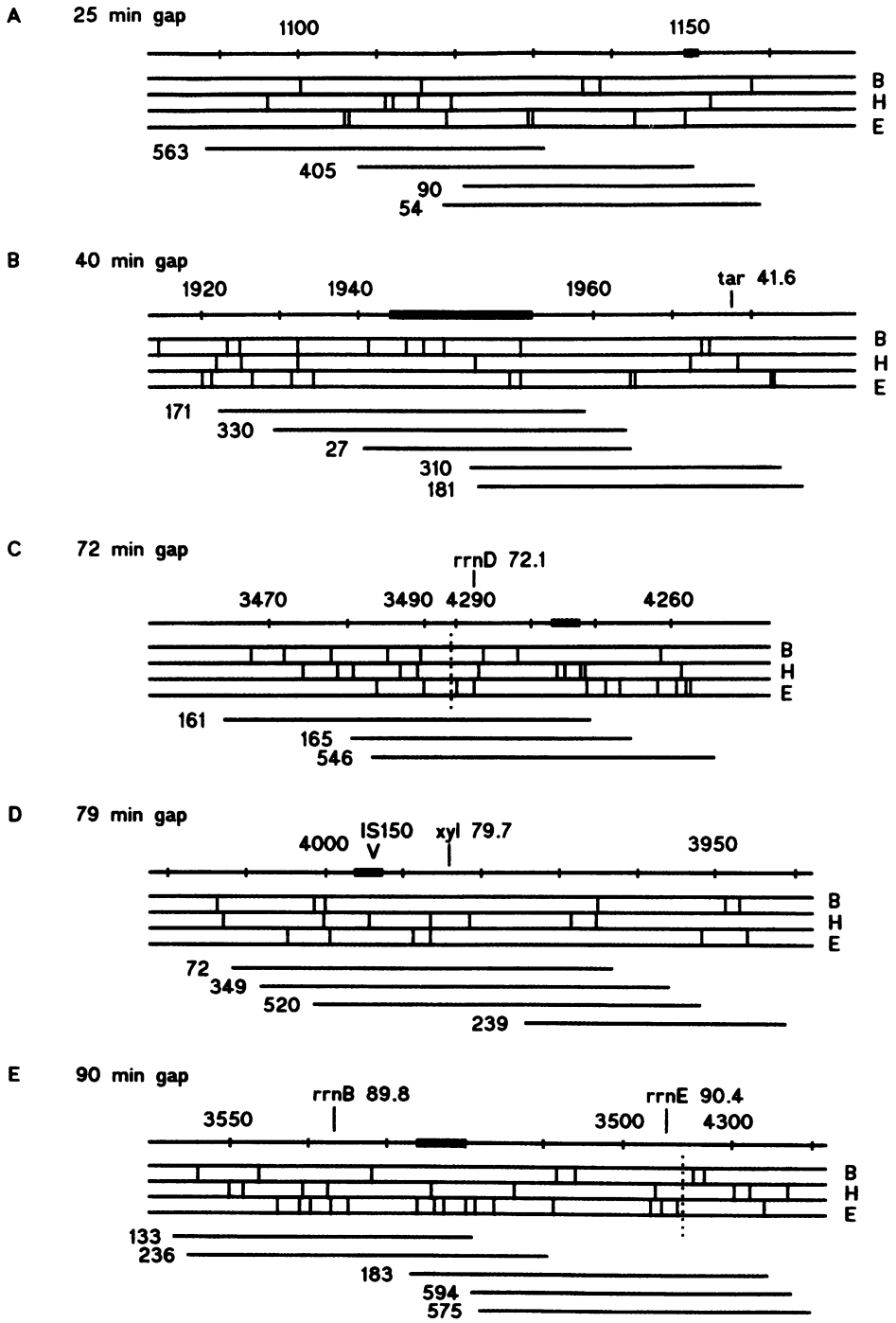


Fig. 2. *Map of the E.coli genome.* Outer circle: Kb-scale (1). Next inner circle: Contigs (black) and gaps (white) in the physical map of strain BHB2600. Encircled numbers: point to gap positions in the Kohara-map in strain W3110. Numbers and letters in innermost circle: loci of cloned genetic markers and rrn loci (see table 1). Dotted bars: inversions in strain BHB2600 (outer) and strain W3110 (inner). Hatched box: triplicate 'casette'-insertion in strain W3110. Chequered box: deletion in strain BHB2600.

W3110 (called briefly Kohara-map) serves in the following as a reference including the map calibration in kb.

Gaps in the map. Our map retains 12 discontinuities (gaps). Some of them are far below CDE size. One of the largest gaps of 40 kb persists at the oriC region (84′), despite efforts to identify suitable CDEs for this map range. In strain K12 803 the corresponding gap has recently been closed (2). Otherwise gaps in our map do not coincide with those remaining in the Kohara-map (see Fig.2) with the exception of one tiny gap at the position of the lpp-gene (1775kb, 36.5′) (45). Fig.3 documents, with restriction patterns for 3 enzymes immediate environments for the five gaps of the Kohara-map which are covered by our map. Thus, besides the mentioned minor exception both maps together complement each other to now form a complete *E.coli* map.

Contig alignment. Discontinuities divide the map into separate segments (higher order contigs) of 60 to 1100 kb in length which cannot be joined per se into a correct map array. In the special case of *E.coli* this may be overcome by a correlation of contig segments


A
25 min gap
1100
1150
B
H
E
563
405
90
54
B
40 min gap
1920
1940
1960
tar 41.6
B
H
E
171
330
27
310
181
C
72 min gap
rrnD 72.1
3470
3490
4290
4260
B
H
E
161
165
546
D
79 min gap
4000
IS150
V
xyl 79.7
3950
B
H
E
72
349
520
239
E
90 min gap
rrnB 89.8
3550
rrnE 90.4
3500
4300
B
H
E
133
236
183
594
575

with cloned gene markers assigned to the classical genetic map (9). For this we have used a collection of map assigned clones (listed in Tab.1), which directly served as DNA probes, for a search in the CDE working set by dot blot tests. The derived contig order in relation to map positions of the marker clones is drawn in Fig.2. Most contigs were found to be adjoined to available marker positions except for two remaining contigs which fit by size without contradiction into corresponding empty map spaces. Thus, all contigs can be considered to be correctly aligned.

Comparison to the map of Kohara et al. (1). Concerning the EcoR1 restriction pattern, there is generally excellent agreement between the Kohara-map and our map except for local deviations and rearrangements observed between strains W3110 and BHB2600. All differences reported below rest on restriction analytical data (for a detailed account see [8]):

1) Differences in restriction sites: As local deviations were scored 15 gains and 13 losses of single restriction sites with conserved additive lengths of the adjoined fragments. These changes are scattered all over the map.

2) Missing or additional restriction fragments are less frequent. The relevant cases are documented in Tab.2 which shows that most of these changes concern single fragments and in only a few cases two or three fragments.

3) Differences in restriction fragment strings: More interesting for evolutional considerations would be pronounced restriction fragment differences extending over several successive fragments. So far, only one possible deviation of this type has been encountered concerning at least five restriction fragments or more at position 1200kb (25′) (see Tab.2). Because of an adjacent gap region this deviation cannot be followed further, so that it cannot be completely ruled out whether this deviation might be part of a small inversion.

4) Insertions, deletions, inversions: Most obvious are extended map rearrangements observed for the two strains. An enormous inversion flanked by the sequence repetitive genes rrnD (3500kb, 72.1′) and rrnE (4300kb, 90.4′) with map discontinuities close to but not identical with each flank was reported for strain W3110 (1, 49). This inversion is obviously not present in our strain BHB2600 (Fig.2 and Fig.3: 72′ and 90′ gap). On the other hand a considerable inversion is seen in strain BHB2600 at map positions 3150kb (64.5′) to 4050kb (77′) (Fig.2). The breakpoints are associated with mutually inverted IS5 insertion elements. Its counterclockwise flank coincides precisely with the IS5 flank of a small but remarkable insertion apparent in strain W3110 as a threefold repeated 'casette' structure stemming from a singular IS5 flanked sequence positioned elsewhere (37). The 'casette' inversion and its singular prototype are absent in our strain BHB2600. Strain BHB2600 also harbours an inversion flanked by mutually inverted IS1 elements (Fig.2)

Fig. 3. *Regions of gaps in strain W3110 mapped in strain BHB2600.* Upper line: callibration scale in kb according to (1). Marker positions in map minutes (9). Solid bars: gap positions. Middle field: restriction patterns for the enzymes BamHI (H), HindIII (H) and EcoRI (E). Lower field: CDEs. Length of bars corresponds to the resolution of the restriction map.

A: Gap length 2 kb. The gap lacks B, H and E sites.

B: Gap length 15 kb. The gap contains some B, H and E sites, which are identical in strains BHB2600 and W3110.

C: Gap length 4 kb. The gap is close to rrnD locus and the associated breakpoint (dotted line) of the inversion of strain W3110.

D: Gap length 3 kb. The gap contains the singular position of IS150, which is present in strains BHB2600 and W3110. Region D within the inversion of strain W3110 is drawn with inverted scale (BHB2600 orientation).

E: Gap length 5 kb. The gap is close to the rrnE locus and the associated breakpoint (dotted line) of the inversion of strain W3110.

Table 2. Differences in the EcoRI restriction pattern between strains BHB2600 and W3110. A, B: Italic bold numbers: absent or additionally present fragments in otherwise unchanged fragment arrays. Brackets: More than one fragment missing at the indicated (=) fragment position. C: At the left side of the array interrupted by agap in strain BHB2600. The underlined 3.3 kb fragment contains the lit gene (25.4').

Map pos.	[kb]	EcoRI pattern
A		Missing fragments in BHB2600:
	830	32, 6.9, ***12,*** 30, 14
	1420	2.9, 0.7, 3.0, ***0.7,*** 1.7, 8.1
	2180	21, 5.9, 3.2, ***1.0,*** 5.3
	2270	12, 1.4, 1.3, 2.6, ***9.8,*** 17, 7.4
	2730	8.8, 2.2, 1.5, ***1.0,*** 5.0
	2770	2.5, 1.0, 0.6, ***(1.2, 1.4, 8.2,)*** = 3.6, 0.8, 2.5
	2950	1.8, 1.1, 11.5, ***5.6,*** 34
	3290	9.2, 16, ***3.5, 6.2, 1.3,*** 6.3 10
	3920	2.1, 25, ***3.7,*** 3.7, 5.2
	4630	26, 15, ***10.5*** = 0.6, 3.3, 11, 3.2
	4700	20, ***(0.5, 18.5,)*** = 12.5, 10.5, 8.7
B		Additional fragments in BHB2600:
	740	5.9, 5.0, 1.7, ***1.4,*** 21, 22
	2050	18, ***1.3,*** 2.4, 2.3, 2.0, 10
	4230	10.6, 4.7, 0.8, ***1.8,*** 4.7
	4040	30, 5.6, ***0.4, 4.2,*** 7.7, 4.8
C		Different pattern between BHB2600 and W3110:
	1200	W3110 .....1.3, 3.7, 3.2, 8.0, 11,
		BHB2600 Gap 15, 1.0, 3.2, 6.5, 3.3, 5.3, 8.8, 17,...

extending from 270kb (5.8') to 330kb (7.1'). This inversion is associated at its clockwise flank with an IS1 dependent deletion of 40kb length.

Insertion elements. In the context of this map construction, and in view of the role of insertion elements in rearrangements, we have also derived for strain BHB2600 complete maps for IS1, IS2, IS3, IS4, IS5, IS30 and IS150. Curiously, the singular IS150 (40) position found for both strains coincides with the small gap at 4000kb (79') of the Kohara-map (Fig.3). The complexities of the IS maps shall be published elsewhere.

DISCUSSION

We have applied cosmids as vectors to physically map the *E.coli* genome. Although others have made use of the considerable insert sizes of cosmids before (1, 2, 3, 4,), they were decidedly disregarded in favour of lambda vectors in previous projects of *E.coli* mapping (1, 3) with the argument that propagation of *E.coli* DNA in *E.coli* might lead, by metabolic interference, to selective losses of cloned material. In fact, in a recent report an incomplete map saturation was in part ascribed to the use of cosmids as vectors (2, 4). Colonies withdrawn from cosmid libraries indeed exhibit a wide growth variability which causes laborious measures to derive a CDE working set of homogeneous quality and high map representation. Our final result shows that with the derived material a close to complete map can well be constructed.

The price paid by use of cosmids is made up by the reduced number of CDEs required to cover the map. 570 CDE units were sufficient to bring our map to the present 95% of saturation leaving 12 small gaps. Thus there might be only a few genomic regions which resist clonal transfer by cosmid vectors. For two of them poor vector maintenance is in

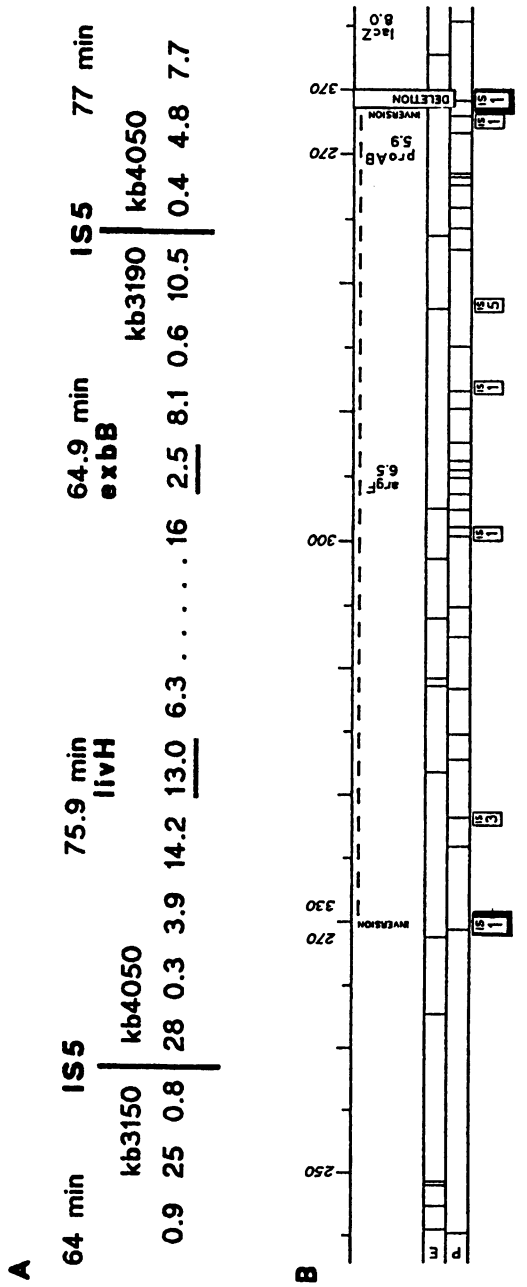
fact indicated: One gap, the last one commonly persisting in the Kohara-map and our map, covers the *lpp gene* (45), a gene which appears to interfere with the maintenance of both cosmid and lambda vectors. The second gap, positioned at the *oriC region* and meanwhile closed in strain K12 803 (2)persists in our map despite determined efforts to close it. OriC carrying cosmids might be lost by interference with the replication initiation mechanism. Or it might be that clustering of Sau3A restriction sites in the oriC region supresses the generation of correct insert sizes (46) a situation which appears to apply for lambda vectors (1). Nothing is known which other gap regions are inaccessible to cosmid dependent mapping. However, at a mapping state leaving a few small gaps, any additional technique or vector may be tried to bridge these gaps.

Concerning our mapping strategy we have chosen dot blot hybridisation as a basic mapping tool because of its reliability to stepwise built up contig interelationships by an all or nothing procedure. Because of the high resolution ($<$ 1kb per contig-overlap) the step width per hybridisation cycle for cosmid derived CDEs is on the average 90 kb (maximally 120kb) per contig. Three complete dot patterns of the CDE working set may be run simultaneously to achieve contig gains of up to 300kb per cycle. In practice, only about 65 (single) hybridisation cycles were required to collect the contig pool for the here presented map including the contig subset covering repetitive elements (IS elements and rrn genes). Instead of using walking procedures (4) we preferred random contig search by use of probes exclusively withdrawn from the pool of yet unidentified CDEs. This provides gradual growth of the contig pool by every hybridisation step and allows to stepwise trace the fusion of low into high order contigs by a genealogy which directly prescribes the map alignment with a resolution down to CDEs. This scheme was only changed when merely smaller gaps were to be filled in the map by systematically selecting probes from gap flanks. In this way 5 gaps were closed.

CDE mapping for contigs $>$ 10 CDEs was overlayed by one enzyme (EcoR1) restriction mapping. Fragments are mapped by two criteria: 1) interruption of an adjoined fragment by a CDE terminus and 2) positioning of a fragment neighboured next to a interrupted fragment in- or outside a CDE element. Thus, one CDE terminus determines the position of one neighboured restriction fragment pair. Towards the end of the mapping process $>$1100 CDE termini for the ordering of maximally 2200 fragments were available, a number sufficient to resolve the about 650 real EcoR1 fragment positions (1) so that only clusters of very small fragments might occasionally escape ordering.

Mapping ambiguities due to the presence of repetitive elements (REs) cause particular problems in hybridisation mapping. This prevalently concerns in the *E.coli* genome the rrn (ribosomal RNA) genes (38) and the variable group of IS (insertion) elements. We have overcome this problem simply by probing the complete CDE working set for the presence of REs by use of probes for rrn (36) and IS 1 (8), 2 (8), 3 (8), 4 (35), 5 (8), 30 (39) and 150 (40) sequences, alltogether a collection of 54 REs distributed in about 100 kb average distance over the *E.coli*-genome. This led to establish a subset of 23 contigs distinguished from each other by internal RE patterns. The RE subset contributes with 44% of map length significantly to the general contig pool. Once available as defined probes, REs are a mapping tool which can help to considerably speed up the mapping process.

The finally derived map for *E.coli* strain BHB2600 may be compared by its array of EcoR1 restriction fragments with the map deviced by Kohara et al. (1) for strain W3110. Both maps retain a few discontinuities which, except for one (lpp), do not coincide so that the maps mutually complement each other to a now complete, clone represented map.


A
64 min
IS5
kb3150
kb4050
0.9 25 0.8 28 0.3 3.9 14.2 13.0 6.3 . . . . . 16 2.5 8.1 0.6 10.5 0.4 4.8 7.7
75.9 min
livH
64.9 min
exbB
kb3190
IS5
kb4050
77 min
B
lacZ
8.0
DELETION
INVERSION
proAB
5.9
argF
6.5
INVERSION
370
270
300
330
270
250
E
P

There is remarkable agreement between the two maps. Expectedly there is a countable number of minor discrepancies concerning gains and losses of restriction sites or whole restriction fragments. Before consistent conclusions can be drawn, it has to be proven in a side by side analysis of the two strains by Southern blot analysis which of the differing positions are true strain dependent deviations and which are due to mapping errors. Of greater interest are rare string deviations concerning several consecutive fragments within an otherwise unchanged environment for which we have found one likely candidate. The question remains by which mechanism such deviations could have been generated.

The best founded strain deviations are extended rearrangements. From their description it is evident that they are exclusively induced by the presence of mutually inverted repetitive elements. This holds for the large inversion flanked by rrnD and rrnE (72′−90′) observed in strain W3110 (1, 49) which is lacking in BHB2600. It also holds for IS flanked inversions found in our strain BHB2600: one IS1 flanked inversion is associated with an obviously IS1 induced deletion (5.8′ to 7.8′) and the other flanked by IS5 (64′−77′) which coincides at one flank with a flank of an insertion made up of a triplicate IS5 associated 'casette' structure uncovered exclusively in strain W3110 (37). Except for induced and selected inversions (42, 49, 50) there are only two reports for other large 'natural' inversions in enterobacterial laboratory strains (41, 51). The fact to straightway have found two yet undescribed inversions in our strain indicates that rearrangements of the *E.coli* genome might be not too rare events. In view of the omnipresence of IS elements (or other repetitive elements) it can be presumed that with time many more or less extended rearrangments will be uncovered within related bacterial strains. On the other hand, it may be asked why the *E.coli* genome, if for instance compared to *Salmonella*, has retained its stability to a still high degree without becoming scrambled within short evolutionary times. This question may be related to the problem of maintenance of the gene dosage gradient along the replicating chromosome (42). It remains open which rearrangements this sensitively balanced machinery tolerates and which not. With an increasing routine for the fast physical mapping of genomes or genome fragments, answers to these questions will be progressively accessible.

Restriction mapping is mostly achieved by 'fingerprint' techniques which detect CDE overlaps by the idendity of restriction patterns (5, 7), while dot blot hybridisation techniques have been so far rarely employed for a primary mapping strategy (4). We have shown that by the latter approach a widespanned contig framework for a cosmid bank in the 4 megabase range can be generated within a relative short time. Manyfold larger bank sizes cannot be treated economically by this technique due to the expansion of dotblot sets with increasing bank sizes. This causes a square increase of the expenditure of work for the prehybridisation part of the procedure. New methods have been recently developed to subdivide large eucaryotic genomes into separable bits from 50 kb up to megabases which can be more conveniently handled in physical mapping procedures (6, 43, 44, 52). Dot blot hybridisation would be an appropriate mapping tool for the large fragment range.

The question remains whether hybridisation techniques are useful for the mapping of

Fig. 4. *Map segments for the inversions in strain BHB 2600.* A: Breakpoints of the inversion between 64′ and 77′. Numbers: EcoR1 fragment sizes in the vicinity around each breakpoint. Vertical bars: IS5 positions at inversion termini (breakpoints). Underlined: probe positions for genetic markers. B: Map segment for the inversion-deletion between 5.8′ and 7.8′. Calibration is according to ref.(1). Horizontal broken line: extent of inversion. E: EcoR1 map segment. P: PstI map segment. Boxes: IS positions (bold: flanking positions).

eucaryotic genomes. In mammalian DNA the overwhelming presence of repetitive retroposon DNA sequences with its broad sequence distribution spectra may cause severe problems. This sequences can be hardly masked by competition procedures, whereas interferences due to loose sequence interrelations of low copy repetitive elements can be sufficiently controlled by hybridisation stringency. Thus, hybridisation techniques may be only applicable for the mapping of genomes of lower eucaryotes. With this restrictions in mind, we are convinced that cosmid supported hybridisation techniques can be applied as an alternative to (or together with) other techniques for the fast and reliable mapping large subgenomic fragments.

ACKNOWLEDGEMENTS

We thank H.I.Trompeter and K.Beginnen for valuable experimental contributions, K.P.Reiners for critical comments and S.Scheff for reading the manuscript. We especially thank Dr. Y. Kohara for providing strain W3110 and three clones of his library, and all those who kindly provided genomic probes.
(Individual clones and a complete restriction map are freely available from R.B. upon request.)

REFERENCES

1. Kohara,Y., Akiyama,K. and Isono,K. (1987) Cell **50**,495–508.
2. Knott,V., Rees,D.J.G. and Cheng,Z., Brownlee,G.G. (1988) Nucl.Acids Res. **16**,2601–2612.
3. Daniels,D.L. and Blattner,F.R. (1987) Nature **325**,831–832.
4. Tabata,S., Higashitani,A., Takanami,M., Akiyama,K., Kohara,Y., Nishimura,A., Yasuda,S. and Hirota,Y. (1989) J.Bac. **171**,1214–1218.
5. Olson,M.V., Duchick,J.E., Graham,M.Y., Brodeur,G.M., Helms,C., Frank,M., MacCollin,M., Scheinman,R. and Frank,T. (1986) Proc.Natl.Acad.Sci.USA **83**,7826–7830.
6. Coulson,A., Waterson,R., Kiff,J., Sulston,J. and Kohara,Y. (1988) Nature **335**,184–186.
7. Coulson,A., Sulston,J., Brenner,S. and Karn,J. (1986) Proc.Natl.Acad. Sci.USA **83**,7821–7825.
8. Birkenbihl,R.P. (1989) Ph.D.thesis, Univ. Cologne., FRG.
9. Bachmann,B.J. (1983) Microbiol.Ref. **47**,180–230.
10. Ohki,M., Tamura,F., Nishimura,S. and Uchida,H. (1986) J.Biol.Chem. **261**,1778–1781.
11. Iwakura,M., Shimura,Y. and Tsuda,K. (1982) J.Biochem. **91**,1205–1212.
12. Chye,M., Guest,J.R. and Pittard,J. (1986) J.Bact. **167**,749–753.
13. Boidol,W., Simonis,M., Töpert,M. and Siewert, G. (1982) Mol.Gen.Genet. **185**,510–512.
14. Brune,M., Schumann,R. and Wittinghofer,F. (1985) Nucl.Acids Res. **13**,7139
15. Hull,E.P., Spencer,M.E., Wood,D. and Guest,J.R. (1983) FEBS Lett. **156**,366–370.
16. Kao,C., Gumbs,E. and Snyder,L. (1987) J.Bact. **169**,1232–1238.
17. Kaiser,K. and Murray,N.E. (1979) Mol.Gen.Genet. **175**,159–174.
18. Kaiser,K. (1980) Mol.Gen.Genet. **179**,547–554.
19. Mechulam,Y., Fayat,G. and Blanquet,S. (1985) J.Bact. **163**,787–791.
20. Stader,J., Matsumura,P., Vacante,D., Dean,G.E. and Macnab,R.M. (1986) J.Bact. **166**,244–252.
21. Kuo,S.C. and Koshland,D.E. (1986) J.Bact. **166**,1007–1012.
22. Bhagwat,A.S., Sohail,A. and Roberts,R.J. (1986) J.Bact. **166**,751–755.
23. Bruni,C.B., Musti,A.M., Frunzio,R. and Blasi,F. (1980) J.Bact. **142**,32–42.
24. Mizuno,T., Chou,M. and Inouye,M., (1983) J.Biol.Chem. **258**,6932–6940.
25. Richaud,F., Richaud,C., Ratet,P. and Patte,J. (1986) J.Bact.**166**,297–300.
26. Parker,J. and Eisenbeis,S.J. (1981) Mol.Gen.Genet. **183**,115–122.
27. Stauffer,G.V., Plamann,M.D. and Stauffer,L.T. (1981) Gene **14**,63–72.
28. Eick-Helmerich,K., Hantke,K. and Braun,V. (1987) Mol.Gen.Genet. **206**,246–251.
29. Macdonald,H. and Cole,J. (1985) Mol.Gen.Genet. **200**,328–334.
30. Nazos,P.M., Antonucci,T.K., Landick,R. and Oxender,D.L. (1986) J.Bact. **166**,565–573.
31. Wek,R.C., Hauser,C.A. and Wesley Hatfield,G. (1985) Nucl.Acids Res. **13**,3995.
32. Rehemtulla,A., Kadam,S.K. and Sanderson,K.E. (1986) J.Bact. **166**,651–657.

33. Little,J.W. (1980) Gene **10**,237–247.
34. Liljeström,P.L. and Liljeström,P. (1987) Nucl.Acids Res. **15**,2213.
35. Klaer,R. and Starlinger,P. (1980) Mol.Gen.Genet. **178**,285–291.
36. Niebel,H. (1987) Diploma thesis, Kiel FRG.
37. Muramatsu,S., Kato,M., Kohara,Y. and Mizuno,T. (1988) Mol.Gen.Genet. **214**,433–438.
38. Boros,I., Kiss,A. and Venetianer,P. (1979) Nucl.Acids Res. **6**,1817–1830.
39. Dalrymple,B., Caspers,P. and Arber,W. (1984) EMBO J. **3**,2145–2149.
40. Schwartz,E., Kröger,M. and Rak,B. (1988) Nucl.Acids Res. **16**,6789–6808.
41. Xia,X-M. and Enomoto,M. (1986) Mol.Gen.Genet. **205**,376–379.
42. Louarn,J.M., Bouché,J.P., Legendre,F., Louarn,J. and Patte,J. (1985) Mol.Gen.Genet. **201**,467–476.
43. Burke,D.T., Carle,G.F. and Olson,M.V. (1987) Science **236**,806–811.
44. Olson,M.V. (1988) Fresenius Z Anal.Chem. **330**,301–302.
45. Nakamura,K. and Inoue,M. (1979) Cell **18**,1109–1117.
46. Sugimoto,K., Oka,A., Sugisaki,H., Takanami,M., Nishimura,A., Yasuda,S. and Hirota,Y. (1979) Proc.Natl.Acad.Sci.USA **76**,575–579.
47. Ish-Horowicz,D. and Burke,J.F. (1981) Nucl.Acids Res. **9**,2989–2998.
48. Chomczynski,P. and Qasba,P.K. (1984) Biochem.Biophys.Res.Comm. **122**,340–344.
49. Hill,C.W., Harnish,B.W. (1981) Proc.Natl.Acad.Sci.USA **78**,7069–7072.
50. Hill,C.W. and Gray,J.A. (1988) Genetics **119**,771–778.
51. Casse,F., Pascal,M. and Chippaux,M. (1973) Mol.Gen.Genet. **124**,253–257.
52. Smith,C.L., Econome,J.G., Schutt,A., Klco,S. and Cantor,C.R. (1987) Science **236**,1448–1453.

This article, submitted on disc, has been automatically converted into this typeset format by the publisher.