Volume 17 Number 15 1989 Nucleic Acids Research

Completion of the detailed restriction map of the *E.coli* genome by the isolation of overlapping cosmid clones

V.Knott, D.J.Blake and G.G.Brownlee

Sir William Dunn School of Pathology, University of Oxford, South Parks Road, Oxford OX1 3RE, UK

Received June 13, 1989; Accepted June 27, 1989

ABSTRACT

Ordered sets of cosmids derived from *E.coli* K-12 803 overlap the 6 remaining gaps left in the physical map of strain W3110 (1,2). We present detailed restriction maps of the gaps and surrounding regions, thus providing a comparison of about 30% of the genome of the two *E.coli* strains. Our analysis shows that there is a high degree of homology between the strains, with only occasional restriction fragment differences. However, the large inversion occurring between *rrnD* (72.1') and *rrnE* (90.4') in strain W3110 is absent in strain 803. Instead, a new inversion and adjacent deletion near *argF* is present in strain 803. The distribution of cosmid clones at, and adjacent to, the gaps shows that all gaps except one were difficult to clone in both λ and cosmid clones. A low copy number cosmid vector, pOU61cos, developed previously (2), was essential for cloning 3 of the 8 gaps.

INTRODUCTION

A description of the exact number and position of all genes within the genome of any organism is clearly a prerequisite for a complete understanding of that organism and its functions. In a small, relatively simple organism, such as *E.coli*, hundreds of genes have been identified by genetic complementation and their positions mapped relative to each other (3). Nevertheless the function of a large proportion of its DNA remains unknown. Searches for new genes would be greatly facilitated if ordered sets of overlapping clones spanning a whole genome were available in conjunction with a detailed restriction map. Ultimately, the sequence of a whole genome is required, and this has been planned for *E.coli*, using a set of ordered clones as starting material for the sequence analysis (4).

Ordered sets of clones can be obtained in different ways. For example, in a *first* approach, libraries of clones can be screened with labelled probes to identify overlapping clones (5,6). By such chromosome walking successive overlapping clones can be isolated, ultimately covering the whole genome. However, this procedure is extremely tedious even if probes for many loci have already been isolated, as is the case for *E.coli*. The presence of repetitive elements is a further complication. A *second* approach is to isolate clones randomly, fingerprint each one and compare the fingerprints to build up sets or 'contigs' of related clones. The random approach would appear to be more appropriate for larger genomes, where less genetic data is available and has been used for example on *Caenorhabditis elegans* (7), on *Saccharomyces cerevisiae* (8) and on *E.coli* (9,2, and present study). The number of clones one has to analyse may be larger than for chromosome walking, as clones have to overlap extensively to obtain a reliable match, but the inbuilt redundancy can help subsequently in the derivation of a restriction map and in the accurate positioning of genetic markers. This method is less subject to interference by repetitive DNA. However, the

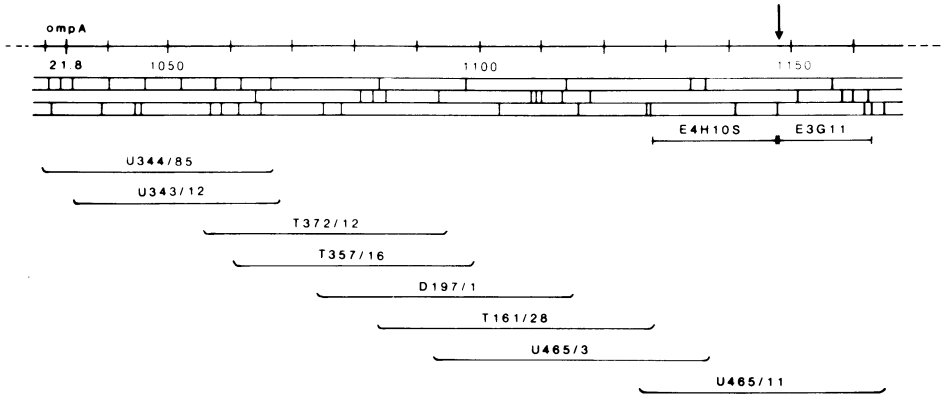


Fig.1. Restriction map of the *ompA* (21.8') region. Cosmid clones are represented by horizontal bars and positioned according to their restriction patterns; their length is approximate. Cosmids derived from different libraries (2) are prefixed by a letter indicating the vector used: C= pDVcos, D= pDV, L= LoristB, S= pDVcosA2, T= pTM, U= pOU61cos. The top band shows the BamHI, the middle band the HindIII and the bottom band the EcoRI restriction map. The map positions are from Kohara et al. (1). The arrow points to the position of the gap in Kohara et al. (1). E3G11 and E4H10S are λ clones (1).

non-random distribution of clones in most libraries proves a serious limitation and makes it extremely difficult to build up large contigs. A *third* approach is to combine the above two procedures and join contigs created by a random screening by chromosome walking. This approach has been used by Kohara et al.(1) in a study of *E.coli* W3110 and a nearly complete set of overlapping λ clones was obtained, leaving only 8 small gaps.

The purpose of this report is to describe a parallel study we have undertaken on *E.coli* K-12 803, using a random approach. Because our study was designed as a model for genomes larger than *E.coli*, we chose to prepare cosmid rather than phage λ libraries. The larger insert size of cosmid clones decreases the number of clones to be handled. We have identified contigs from our ordered cosmid library that overlap all the gaps left in the map of Kohara et al. (1), proving that the entire genome of *E.coli* can be cloned in either cosmid or λ clones. A comparison of the restriction map of about 30% of the

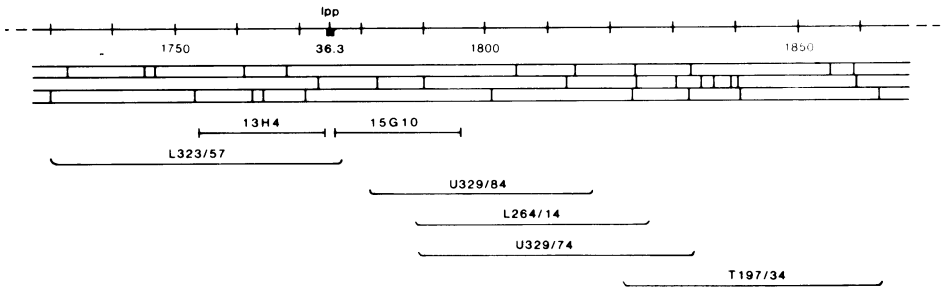


Fig.2. Restriction map of the *lpp* (36.3') region. Scale and symbols are as in Fig.1. The solid rectangle shows the position of the gap in Kohara et al. (1). 13H4 and 15G10 are λ clones (1).

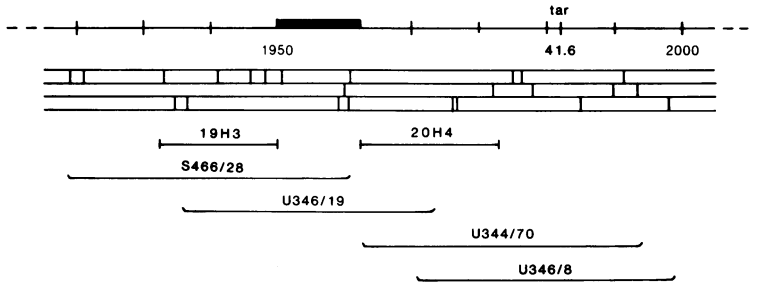


Fig.3. Restriction map of the *tar* (41.6') region. Scale and symbols are as in Figs.1&2. Note that the scale is discontinuous over the gap. The scale in Kohara et al. (1) allows only 2 kb for the gap, whereas it is actually 12 kb long. 19H3 and 20H4 are λ clones (1).

genome of strains 803 and W3110 of *E.coli* emerges from this study. We present further evidence for the usefulness of a low copy number cosmid vector, pOU61cos, developed previously (2), for the cloning of 'difficult' sequences.

MATERIALS AND METHODS

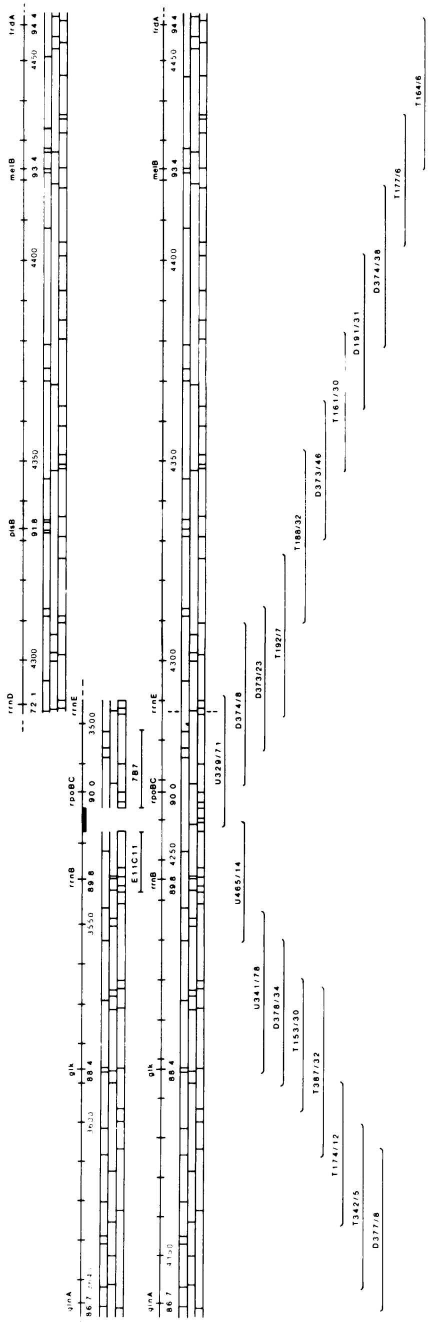
The strain mapped is *E.coli* K-12 803 [*hsdS* (*rk*$^-$, *mk*$^-$) *gal*$^-$, *met*$^-$] (10). The construction of the cosmid libraries, the HinfI fingerprinting and restriction mapping of cosmids are all described in (2). Southern blots were carried out using standard methods.

Screening of the E.coli 803 cosmid library in the vector pOU61cos.

E.coli NM554 [*sup*°, *hsdR, recA*13, (11)] infected with the *E.coli* cosmids were plated on Hybond-C extra membranes (Amersham) at low density and grown at 30°C for 20–24 hours on L-plates containing 30 μg/ml ampicillin. At this stage the colonies were barely visible. Duplicate replicas were made onto Hybond-C membranes by carefully placing the duplicate membranes (prewetted by placing on sterile L-plates) onto the 'master' membrane laid on moist sterile 3MM paper. Both masters and replicas were then grown for 18 hours at 30°C (or until the colonies were clearly visible). The master plates were then stored at 4°C in sealed plastic bags. The replicas were heat induced at 45°C for 4 hours. Then the colonies were lysed, the DNA fixed onto the dry membranes by UV irradiation and the filters pre-hybridised and hybridised following the manufacturer's instructions. Positively hybridising colonies were subsequently picked from the master plate and grown as in (2).

RESULTS

Previously (2), we described the generation of a database of over 2000 ordered cosmid clones of *E.coli* K-12 803. Briefly, these clones were grouped into 58 contigs ranging in size from 40 to 300 kb and were estimated to represent 90% of the *E.coli* genome. We therefore hoped to find clones in our database which would overlap the 8 gaps left in the map of Kohara et al.(1). We had previously established the position of 19 of our contigs on the *E.coli* genetic map (3) and two of those overlapped the gaps at *oriC* and *pyrB* (2). The simplest way to identify clones that overlapped the remaining 6 gaps was to HinfI fingerprint the λ clones flanking the gaps (kindly provided by Y.Kohara) and match them against our database. In this way we found contigs which overlapped gaps


glnA
glk
rrnB
rpoBC
rrnE
rrnD
plsB
melB
frdA
86.7
88.4
89.8
90.0
72.1
91.8
93.4
94.4
3500
3550
4250
4300
4350
4400
4450
7B7
E11C11
U329/71
U465/14
D374/8
D373/23
T192/7
T188/32
D373/46
T161/30
D191/31
D374/38
T177/6
T164/6
U341/78
D378/34
T153/30
T387/32
T174/12
T342/5
D377/8

2 (*lpp*), 3 (near *tar*), part of 4 (*rrnE*), 6 (*xyl*), and 7 (*rrnD*). The HinfI fingerprint from the two λ clones flanking gap 1 (near *ompA*) and the λ clone flanking one side of gap 4 (*rrnE*) were uninformative, i.e. they showed too few restriction fragments in our fingerprinting system to give a reliable match with cosmid clones from the database (data not shown). Cosmid clones overlapping these 2 gaps were found by chromosome walking (see Materials and Methods), using restriction fragments derived from the λ clones to probe a library of low copy number cosmid clones prepared with the vector pOU61cos. HinfI fingerprints of the cosmid clones isolated in this way were subsequently matched against the database in order to identify related contigs. Restriction maps of our cosmid contigs with BamHI, HindIII and EcoRI (2) aligned to the published restriction map of *E.coli* W3110 (1) allowed us to deduce the length of the gaps and the restriction sites within the gaps. The same procedure also allowed us to make a detailed comparison of the physical maps of over 30% of strains 803 and W3110. We further present (see below) histograms of the distribution of cosmid clones around the gaps which indicate that the same regions that presented cloning difficulties in λ (1) also presented difficulties in cosmid cloning. Assuming the average cosmid insert length to be 40 kb, the 4700 kb of the *E.coli* genome can be represented by 120 cosmids end to end. There are 2195 ordered cosmids in our database, and 444 of them derive from the low copy number library pOU61cos. Therefore, assuming a random distribution, any given locus should be present in about 18 cosmids, 4 of them in the pOU61cos vector.

Detailed information on the clones covering the gaps and the restriction map which we established within and adjacent to these gaps is now presented—except for gaps 5 and 8 (*oriC* and *pyrB*), for which this data has already been published (2). We also present a map covering the genetic distance from 5′−9′ showing a deletion and inversion around *argF* in strain 803.

Gap 1 (near ompA 21.8′)

There is no gap here, as we deduce that the two flanking λ clones abutt (see Fig.1). Probing a pOU61cos library with the two λ clones adjoining the gap, E4H10S and E3G11 (see Materials and Methods), allowed the isolation of 2 related cosmid clones, but only one of these overlaps the gap region (see Fig.1). HinfI profiles of these two cosmids were matched against the database (2) and identified a contig which extended from clone U465/3 to *ompA* at 21.8′ (see Fig.1). No contig could be identified that overlapped clone U465/11 and extended further to the right, perhaps because this cosmid has only 4 HinfI sites that label in our fingerprinting system (2). There is no difference in the restriction map between the strains W3110 and 803 over the 130 kb overlapped by cosmid clones (see Fig.1). However, the clone frequency is generally low and the gap and nearby clones are mainly from the pOU61cos library (see Fig.8A).

Gap 2 (lpp 36.3′)

This gap is covered by clone L323/57, which, however, fails to overlap with clone U329/84 (Fig.2). On Southern blots, L323/57 shows a 3 kb EcoRI (insert+vector), 7 kb HindIII (insert+vector) and a 14 kb BamHI (insert+vector) band hybridising with the 15G10 λ clone. The size of the restriction fragments crossing the gap (see Fig.2) was confirmed

Fig.4. Restriction map (bottom) of the region from *glnA* (86.7′) to *frdA* (94.6′). Scale and symbols are as in Figs.1&2. The restriction maps from one side of the *rrnD* and *rrnE* locus in strain W3110 (1) are also shown (top and middle, respectively). The deduced inversion break point is marked by a vertical dotted line. E11C11 and 7B7 are λ clones (1).

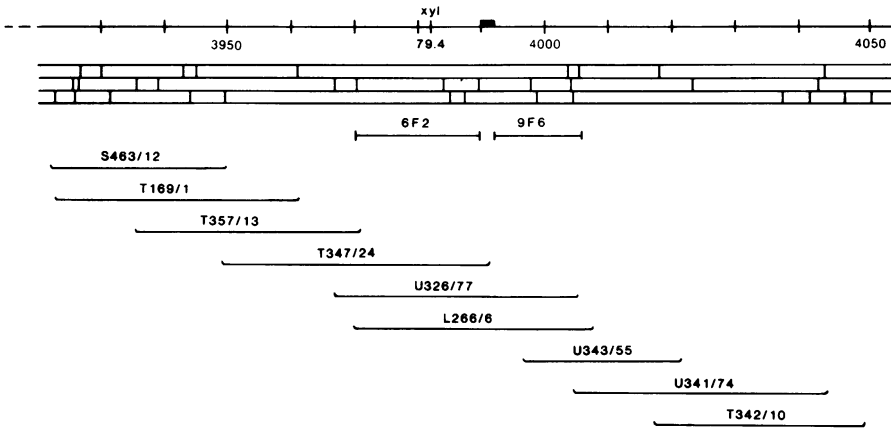


Fig.5. Restriction map of the *xyl* (79.4') region. Scale and symbols are as in Figs.1&2. 6F2 and 9F6 are λ clones (1).

by Southern blot analysis of restricted genomic 803 DNA (data not shown). We deduce that the gap is less than 1 kb long. In strain 803 there is no BamHI site to the right of *lpp* at map position 1782 (in disagreement with Nakamura et al., 12), as indeed there is no BamHI site in the λ clone 15G10. At map position 1830, strain 803 has an additional HindIII site, creating a 4.2 kb and 2 kb fragment instead of the 6 kb fragment described in strain W3110. The frequency of clones over this region is lower than average (see Fig.8B)

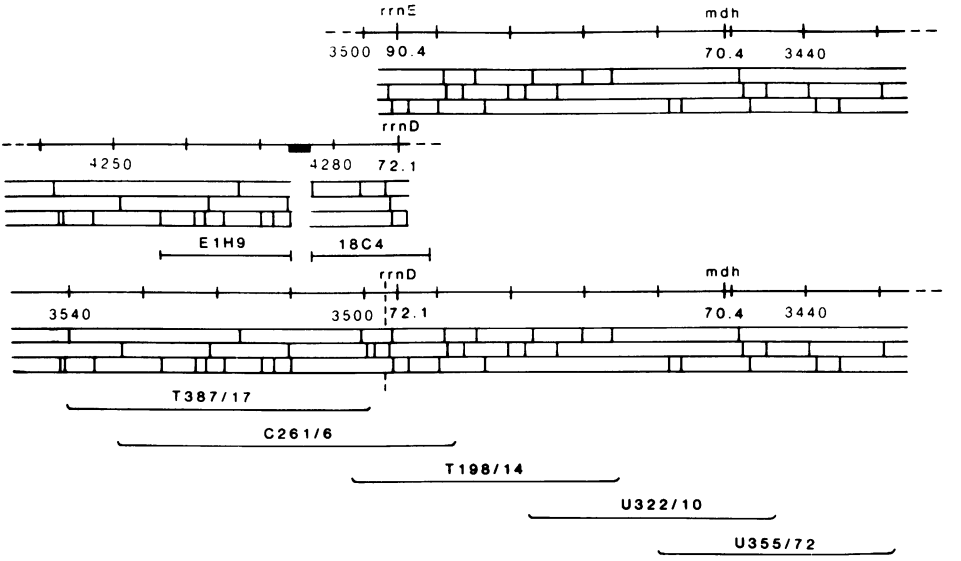


Fig.6. Restriction map of the *rrnD* (72.1') region (bottom). Scale and symbols are as in Figs.1&2. The restriction maps of the *rrnD* and *rrnE* regions in (1) are also shown (middle and top, respectively). The deduced inversion break point is marked by a vertical dotted line. E1H9 and 18C4 are λ clones (1).

and they are mainly from the low copy number cosmid library around *lpp*. Around map position 1850 they are mainly from the higher copy number cosmid libraries.

Gap 3 (near tar 41.6')

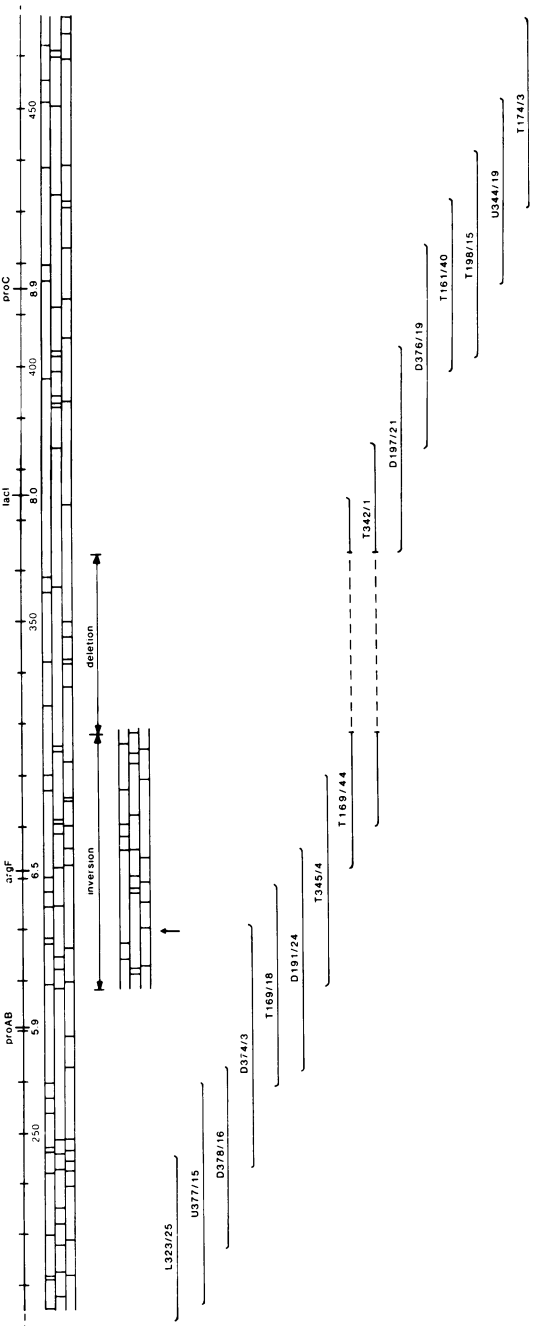
Gap 3 is about 12 kb long and covered by a single clone, U346/19 (see Fig.3). Another clone, S466/28, partially covers the gap. The restriction map for BamHI, HindIII and EcoRI in the gap region was confirmed by Southern blotting of restricted *E.coli* 803 DNA, probed both with the 1.7 kb EcoRI fragment of U346/19 (see Fig.3) and with each of the flanking λ clones. There are several differences between the maps for strain 803 and W3110 in this region. To the right of tar the presence of an additional 3.3 kb HindIII fragment as well as the absence of several EcoRI fragments indicate a rearrangement around map position 1990. There is also a difference in BamHI, HindIII and EcoRI patterns around map position 1920. However this region is only covered by a single clone both in our map and in the Kohara map and the data is therefore intrinsically less reliable. The clones shown in Fig.3 are the only ones from our database covering this region and 3 out of 4 are from the pOU61cos library.

Gap 4 (rrnE 90.4')

Clone U329/71 nearly covers the gap of 7 kb, missing about 1 kb (see Fig.4). No contig could be identified within our database matching the λ clone E11C11 as its HinfI profile gave only 2 bands. Therefore a low copy number cosmid library was screened with the 9 kb EcoRI fragment from E11C11 and clone U465/14 was isolated (see Materials and Methods). It's HinfI fingerprint was then matched against the database, identifying the contig shown to the left of the gap (see Fig.4). Note that the relatively short overlap (less than 10 kb) between U465/14 and U341/78 contained sufficient HinfI bands for a successful match. The short overlap (about 2 kb) between cosmids U329/71 and U465/14 has been confirmed by a Southern blot probed with the 13 kb HindIII fragment of U465/14 (data not shown). Since strain 803 does not contain the inversion between *rrnD* and *rrnE* of strain W3110 used by Kohara et al. (1), a comparison of the restriction maps of the two strains defines the end points of this inversion (see also gap 7, below). We deduce that it lies within the 1.7 kb EcoRI fragment left of *rrnE* at map position 4287 (see Fig.4). Over the 320 kb covered by the overlapping cosmids in Fig. 4 there are only 2 small differences between the strains: at map position 4335 a BamHI site is missing in strain 803, resulting in a larger fragment, and near map position 4430 two BamHI sites are missing. We deduce that clone D377/8 near *glnA* overlaps with clone T387/24 of the contig covering the *oriC* region because they share a 5 kb BamHI fragment (2, and data not shown). This overlap could not be identified from the HinfI fingerprint. The clone frequency is low around the gap region and to the left of it (see Fig.8C), where it is only represented in pOU61cos clones. However, the region left of *glk* at map position 3600 is over-represented: the master clone T174/12 contains 50 near-identical clones, more than twice the predicted number. Similarly the region around *melB* at 93.4' contains more clones than expected: 35 near-identities are contained within the master clone T177/6.

Gap 6 (xyl 79.4')

This gap is 2−3 kb long and is covered by 3 clones from our database, 2 of them from the pOU61cos library (see Fig.5). There is only one difference between the strains over the 130 kb compared: strain 803 contains an additional HindIII site at 4023. The clone frequency is low around the gap and this region contains a high proportion of clones from pOU61cos (Fig.9A). To the left of the gap around map position 3950 the clone frequency is higher than expected: 54 identical or near identical cosmid clones were identified and


proAB
5.9
argF
6.5
lacI
8.0
proC
8.9
250
350
400
450
inversion
deletion
L323/25
U377/15
D378/16
D374/3
T169/18
D191/24
T345/4
T169/44
T342/1
D197/21
D376/19
T161/40
T198/15
U344/19
T174/3

mapped within either T169/1 or T357/13, whereas one would expect only about 23 clones covering this length of DNA in a random clone distribution [this was calculated by dividing the total number of ordered cosmids in the database (2195) by the fraction of the total *E.coli* genome represented by the 2 master clones (50 kb)].

Gap 7 (rrnD 72.1')

The gap is 2−3 kb long and is covered by 2 cosmid clones neither of them from the low copy number library (see Fig.6). We deduce from a comparison of the restriction maps in strain 803 and W3110 that the break point of the inversion is about 1 kb left of the 2.2 EcoRI of the *rrnD* locus (see Fig.6). As a result, the 4.5 kb BamHI fragment at map position 3500 in strain 803 is only 3.3 kb long in W3110 and in the λ clone 18C4. A Southern blot of a BamHI digest of the cosmid clones and strain 803 DNA probed with the 3.3 kb BamHI fragment of 18C4 confirmed this (data not shown). There are three minor differences between the restriction maps of strains 803 and W3110. There is a BamHI site immediately to the right of the break point at *rrnD* in strain 803, and two additional HindIII sites in strain 803 about 2 kb left of the break point of the inversion. The exact position of the two HindIII restriction fragments within the 4.5 kb BamHI fragment at *rrnD* has not been determined. The overall clone frequency is low, although the number of low copy number cosmid clones is close to the theoretical value for the length of DNA under consideration (Fig.9B).

Contig 33 (5'−9')

This large contig shows an inversion in the *argF* region flanked by a deletion between *argF* and *lacI* in strain 803 (see Fig.7). The end points of the inversion and deletion have not been accurately defined as the analysis is complicated by the fact that they are adjacent. Apart from this large rearrangement, there are only 2 strain differences. A small EcoRI fragment at map position 316 is missing in strain 803 (see inverted restriction map below at map position 292), and the BamHI pattern at map position 455 is different in strain 803. The latter, however, may be artefactual as it is present in the single cosmid clone, T174/3, only. We have established the approximate position of a NotI site within the 15 kb HindIII fragment near *argF* (see Fig.7). The clone frequency (Fig.9C) shows that lacI is preferentially cloned in cosmids as there are nearly twice as many clones in the T169/44 and T342/1 master clones than are expected in a random distribution. To the right of *proC* there are very few clones and most are from the pOU61cos library.

DISCUSSION

In overlapping the 8 gaps left in the clone map of *E.coli* (1) with contigs from our cosmid database (2, and this paper) we have shown that the entire *E.coli* genome is clonable in either cosmid or λ clones. Kohara et al. (1) deduced that the genome was 4704 kb long; by adding the DNA contained in the gaps this figure now becomes 4736 kb. Out of the 8 gaps, only one (gap8, *pyrB*, see ref. 2) was bridged by the number of clones expected if one assumes a random distribution of clones in our database (see Results above). The other gaps were overlapped by single clones only, except for gap6 (*xyl*), which was overlapped by 3 clones (two of them from the low copy number library pOU61cos), and gap 7 (*rrnD*), which was overlapped by two clones. Furthermore the clone distribution

Fig.7. Restriction map of the 5'−9' region, with the inverted and deleted section in strain 803 (280 kb−365 kb, below), compared to the corresponding section in strain W3110 (above). Scale and symbols are as in Figs.1&2. The end points of the inversion and deletion are approximate (see text). The arrow indicates the approximate position of a NotI site.

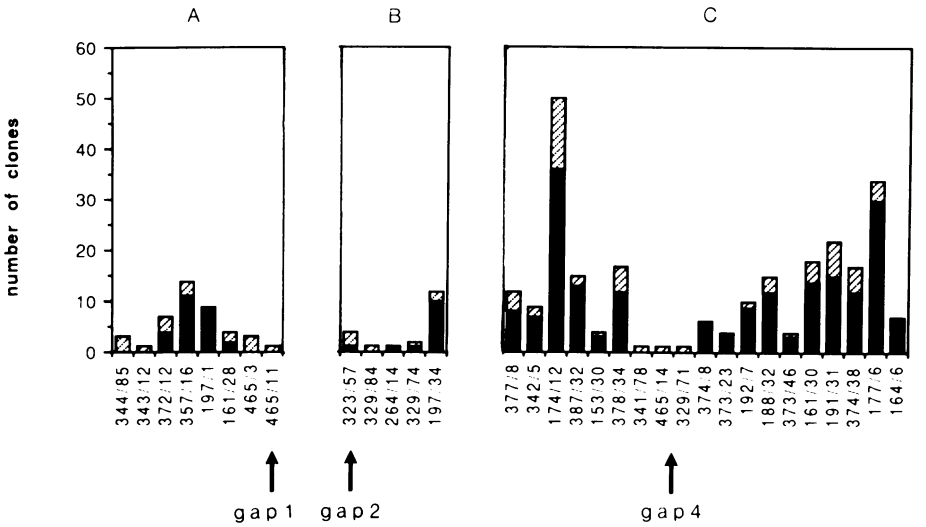


Fig.8. Histograms A, B, and C, of the frequency of cosmids around the gaps 1, 2 and 4, respectively. Each histogram shows the number of near-identical clones contained within each master clone shown on Figs.1, 2 and 4. Note that clones 465/3, 465/11 and 465/14 were not picked randomly (see Results). The position of the gaps is indicated by an arrow. Striped blocks = clones from the low copy number library pOU61cos, solid blocks = others.

in the regions adjoining the gaps is low for all except gap 8 (see ref. 2, and Figs. 3, 8 and 9) and composed mostly of low copy number cosmids. This indicates that the same regions that were difficult to clone in λ (1) were also difficult to clone in cosmids.

The use of the low copy number vector pOU61cos was essential for the cloning of 3 of the 8 gaps, and two further gaps were overlapped by clones in the LoristB vector (13), which has a higher but controlled copy number (see Figs. 3, 8&9). This shows that cloning at low copy number can overcome the adverse effects of gene dosage. However, there may be additional reasons for cloning difficulties. For example, the presence of strong promoters in the cloned insert could interfere with the replication of the cosmid or λ clone. The inclusion of suitable terminator sequences might alleviate this problem (14). When cloning exogenous DNA in *E.coli*, losses of sequences containing repeats or methylated DNA can occur; these can be minimised by a judicial choice of host strain (11, 15).

The restriction maps described here and in (2) allow a comparison of 30% of the genome of strain 803 with that of strain W3110 (1). Apart from minor differences (i.e. the presence or absence of single restriction sites) which can be explained by point mutations or minor mapping errors, we find that the two strains are remarkably similar. A large inversion betwee *rrnD* (72.1′) and *rrnE* (90.4′) in strain W3110 has already been noted and studied (16). Strain 803 lacks this inversion and its end points can be quite accurately defined by comparing the restriction maps of the two strains (Figs.4&6). We have also discovered a 40 kb inversion flanked by a 30 kb deletion between *proAB* and *lacI* in strain 803 (Fig.7). Several insertion sequences have been reported in this region and such sequences are known to promote inversions and deletions (17). There are also several shorter regions where

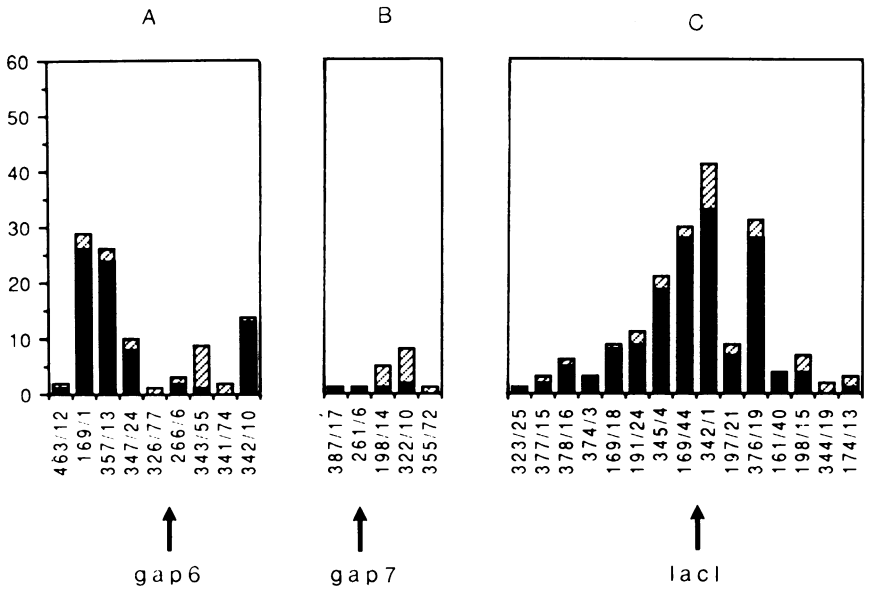


Fig.9. Histograms A, B and C, of the frequency of cosmids around the gaps 6,7 and in the 5–9′ (*lacI*) region respectively. For identification of blocks and symbols see Fig.8.

the restriction maps differ between the two strains (see Results), indicating some minor rearrangement of the DNA between the two strains. Studies of such regions will be facilitated by the availability of ordered sets of clones from both strains. A comparison of restriction maps derived from ordered sets of clones from different strains and different species may be an efficient way to establish their phylogeny and at the same time provide the starting material for comparative studies of common functions.

We undertook this study as a model to test random mapping of large stretches of DNA, ultimately hoping it might be applicable to the whole human genome. *E.coli* seemed an ideal choice because of its well characterised genetic map (3) and moderate genome size (about 4700 kb). As some problems were foreseeable in growing *E.coli* sequences in *E.coli* due to gene dosage (notably genes involved in replication), we also wanted to investigate ways of obtaining more representative libraries. The distribution of cosmids in the different contigs spanning most of the *E.coli* genome gives a better representation of the overall randomness of a library than the frequency for any given locus. Thus we can show that the low copy number cosmid vector pOU61cos (2) allowed us to clone regions difficult to clone even in λ and gave a more random distribution of clones overall than was obtainable using higher copy number cosmids (Figs.8&9). No doubt the cloning strategies and mapping methods used here will need to be refined and adapted for the mapping of larger genomes, notably by the inclusion of yeast artificial chromosomes (YAC) as cloning vectors (18,19). Our experience shows that a random approach, alone, is an impracticable strategy for the mapping of large stretches of DNA and will have to be complemented by chromosome walking or YAC-walking to achieve completion.

ACKNOWLEDGEMENTS

We thank the Wellcome Trust for financial support and Dr.Y.Kohara for providing the λ clones. D.J.B. is supported by a S.E.R.C. studentship.

REFERENCES

1. Kohara,Y., Akiyama,K. and Isono,K. (1987). Cell **50**, 495–508.
2. Knott,V., Rees,D.J.G., Cheng,Z. and Brownlee,G.G. (1988). Nucl.Acids Res. **16**, 2601–2612.
3. Bachmann,B.J. (1983). Microbiol.Rev. **47**, 180–230.
4. Anderson,A. (1989). Nature (London) **338**, 238.
5. Clarke,L. and J. Carbon. (1976). Cell **9**, 91–99.
6. Tabata,S., Higashitani,A., Takanami,M., Akiyama,K., Kohara,Y., Nishimura,Y., Nishimura,A., Yasuda,S. and Hirota,Y. (1989). J.Bacteriol. **171**, 1214–1218.
7. Coulson,A., Sulston,J., Brenner,S. and Karn,J. (1986). Proc.Natl.Acad.Sci.USA **83**, 7821–7825.
8. Olson,M.V., Dutchik,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.
9. Daniels,D.L. and Blattner,F.R. (1987). Nature (London) **325**, 831–832.
10. Wood,W.B. (1966). J.Mol.Biol. **16**, 118–133.
11. Raleigh,E.A., Murray,N.E., Revel,H., Blumenthal,R.M., Westaway,D., Reith,A.D., Rigby,P.W.J., Elhai,J. and Hanahan,D. (1988). Nucl.Acids Res. **16**, 1563–1575.
12. Nakamura,K., Katz-Wurtzel,E.T., Pirtle,R.M. and Inoue,M. (1979).J.Bacteriol. **138**, 715–720.
13. Little,P.F.R. and Cross,S.H. (1985). Proc.Natl.Acad.Sci.USA **82**, 3159–3163.
14. Gibson,T.J., Coulson,A.R., Sulston,J.E. and Little,P.F.R. (1987). Gene **53**, 275–281.
15. Ishiura,M., Hazumi,N., Koide,T., Uchida,T. and Okada,Y. (1989). J.Bacteriol. **171**, 1068–1074.
16. Hill,C.W. and Harnish,B.W. (1981) Proc.Natl.Acad.Sci.USA **78**, 7069–7072.
17. Savič,D.J., Romac,S.P. and Ehrlich,S.D. (1983). J.Bacteriol. **155**, 943–946.
18. Burke,D.T., Carle,G.F. and Olson,M.V. (1987). Science **236**, 806–812.
19. Coulson,A., Waterston,R., Kiff,J., Sulston,J. and Kohara,Y. (1988). Nature (London) **335**, 184–186.

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