Nucleic Acids Research, Vol. 20, No. 5 ***1083–1085***

Stable propagation of cosmid sized human DNA inserts in an F factor based vector

Ung-Jin Kim, Hiroaki Shizuya, Pieter J.de Jong[1], Bruce Birren and Melvin I.Simon*
Division of Biology, 147-75, California Institute of Technology, Pasadena, CA 91125 and [1]Human Genome Center, L-452, Lawrence Livermore National Laboratory, Livermore, CA 94550, USA

Received December 2, 1991; Accepted January 5, 1992

ABSTRACT

Instability of complex mammalian genomic DNA inserts is commonplace in cosmid libraries constructed in conventional multicopy vectors. To develop a means to construct stable libraries, we have developed a low copy number cosmid vector based on the *E.coli* F factor replicon (Fosmid). We have tested relative stability of human DNA inserts in Fosmids and in two conventional multicopy vectors (Lawrist 16 and Supercos) by comparing the frequency of changes in restriction patterns of the inserts after propagating randomly picked human genomic clones based on these vectors. We found that the clones based on Fosmid vector undergo detectable changes at a greatly reduced frequency. We also observed that sequences that undergo drastic rearrangements and deletions during propagation in a conventional vector were stably propagated when recloned as Fosmids. The results indicate that Fosmid system may be useful for constructing stable libraries from complex genomes.

INTRODUCTION

Due to the presence of multiple repeated elements and modified bases, mammalian genomic DNA cloned in *E. coli* frequently undergoes modifications including rearrangements and deletions. Although the instability of human DNA inserts in genomic cosmid libraries is reported (1) and widely acknowledged in anecdotal fashion, quantitative analysis on the instability of complex inserts in genomic libraries has not been documented to date. Use of host strains with various deficiencies in the recombination system could improve the stability of mammalian genomic libraries (2–5). In addition to the major host recombination functions, DNA repair activities and normally minor pathways that mediate recombination in the the presence of DNA fragments containing highly reiterated mammalian DNA sequences in multiple copies may be another factor causing frequent rearrangement and deletion in *E. coli*. These host systems cannot all be eliminated since the viability and growth rate must be considered.

Therefore we have developed a new cosmid cloning system (referred to as Fosmid system) that employs the use of a low copy number plasmid based on the F factor replication origin. We have demonstrated the ability of this vector to stably propagate human genomic DNA inserts in comparison with that of other conventional multicopy cosmids.

MATERIALS AND METHODS

Construction of pFOS1

Standard procedures for recombinant DNA experiments were used (6). The plasmid pFOS1 was constructed by fusing pBAC (Shizuya et al., in preparation), an F replicon based plasmid with lambda *cosN* site, to pUCcos, a pUC derivative containing *cosN* site, by homologous recombination in *E. coli* through the shared *cosN* site (Figure 1). pBAC and pUCcos were transformed to the strain C600r$^-$m$^+$ sequentially, and the transformants grown on LB plates containing ampicillin and chloramphenicol were picked. Miniprep plasmid DNA prepared from the transformant containing fused replicons was diluted and transformed to the strain D110polA$^-$, in which pUC replication origin cannot function and therefore only the plasmids containing the F factor replication origin can survive. Colonies resistant to both ampicillin and choloramphenicol were picked, and screened by miniprep and restriction digestion. pFOS1 DNA thus obtained was transferred to the pop2136 strain (7) for large quantity preparation.

Construction of genomic libraries

Construction of Fosmid, Lawrist 16 (P.de Jong, unpublished), and Supercos (Stratagene) libraries was performed according to the procedure for double *cos* cosmids (8). To generate two arms, the plasmids were completely digested with AatII (pFOS1 and Lawrist 16) or XbaI (Supercos), dephosphorylated by using calf intestine alkaline phosphatase (BRL), and digested with BamHI. The arms were ligated to human DNA isolated from the leukemic cell line Molt4 and partially digested with MboI followed by dephosphorylation using HK phosphatase (Epicentre). The ligated DNA was *in vitro* packaged by using Gigapak Gold packaging system (Stratagene), the cosmid particles were transfected to *E.coli* strain DH5MCR (BRL), and the cells were spread onto appropriate selective media.

* To whom correspondence should be addressed

Propagation and analysis of the clones

18 random clones were picked from each of the libraries and grown in 2 ml LB with appropriate antibiotics at 37°C overnight. These were referred to as primiary cultures and considered to be 0 generation. Cells from these primary cultures were propagated serially by diluting 10^6 fold daily and growing overnight at 37°C. Each passage was considered to represent about 20 generations. 1.5 ml aliquots of the cultures were removed periodically, and the cosmid DNAs were prepared by alkaline miniprep procedure (6). Approximately 1/5 (in case of Fosmids) or 1/20 to 1/50 (in case of cosmids) of the miniprep DNAs were digested with restriction enzymes and loaded onto single wells of agarose gels.

Recloning of cosmid inserts to Fosmid vector

Cosmid DNAs from primary cultures of Supercos clones were digested with NotI to excise the insert DNA. These were ligated to Fosmid arms generated by digesting pFOS1 at NotI sites flanking BamHI/HindIII cloning sites (Shizuya et al, in preparation) after AatII digestion and dephosphorylation. The ligated DNAs were *in vitro* packaged and transfected to DH5αMCR.

RESULTS AND DISCUSSIONS

As shown in Figure 1, pFOS1 is a bireplicon plasmid, and exists in high copy number in *E.coli* due to the pUC replication origin. However, the pUC-derived portion of the plasmid is removed during *in vitro* packaging after ligating the arms to insert DNA (see Materials and Methods), rendering the vector with insert under the control of F factor replication origin.

We have compared the relative stability of human DNA cloned in pFOS vector with that of DNA cloned in two conventional multicopy vectors, Lawrist 16 and Supercos, which are based on bacteriophage lambda and pBR replication origins, respectively. Human genomic libraries based on three different vectors were prepared in parallel according to the procedure developed for double *cos* cosmids (8). Fosmid and cosmid clones

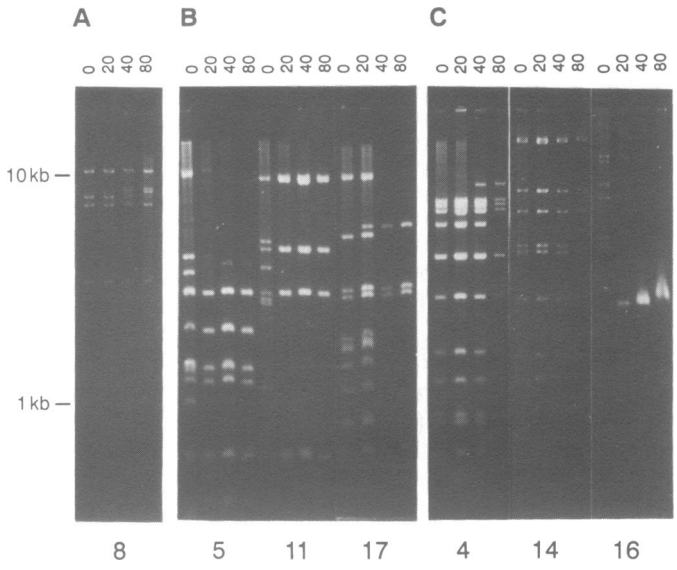

Figure 2. Ethidium bromide stained agarose gels showing EcoRI fragments of miniprep DNA's from unstable Fosmid (A), Lawrist 16 (B), and Super cos (C) clones. Each of the clones were grown to 0, 20, 40, and 80 generations, minipreped, digested with EcoRI and run in 0.7% agarose gels. Clone numbers are shown under the panels.

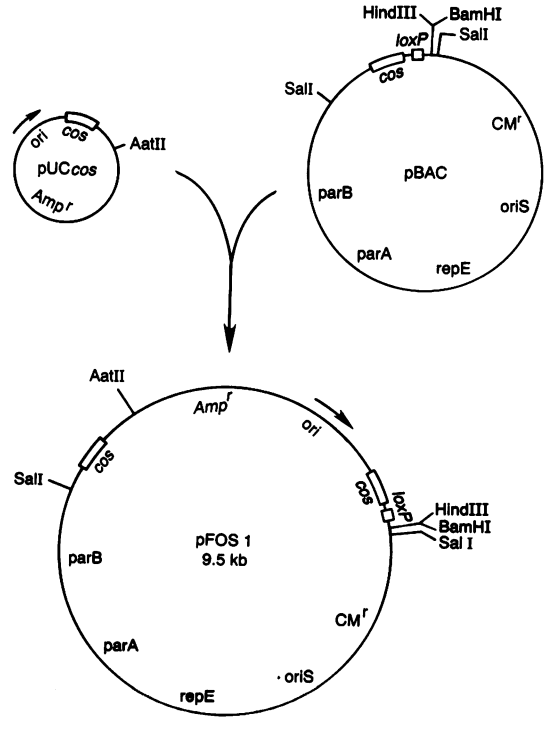

Figure 1. Construction of pFOS1. Details of the procedures are described in MATERIALS AND METHODS. Two NotI sites directly flanking BamHI/HindIII cloning sites are not shown. Amp^r and CM^r indicate resistances to ampicillin and chloramphenicol, respectively.

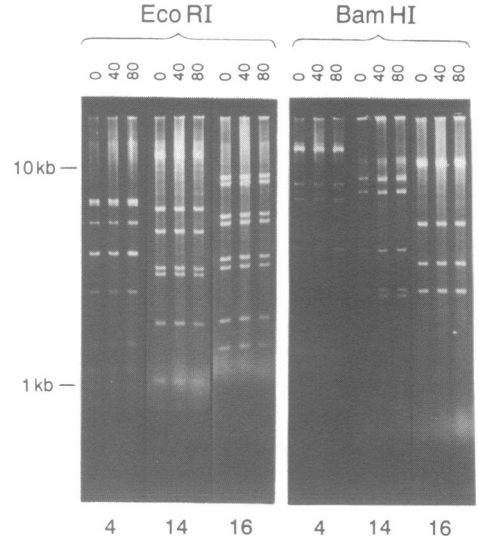

Figure 3. Ethidium bromide stained agarose gel showing EcoRI and BamHI fragments of miniprep DNA's prepared from Fosmid clones containing human DNA inserts excised from unstable Supercos clones #4, 14, and 16 (Figure 2C) after 0, 40, and 80 generations.

were randomly picked, grown serially to multiple generations, and DNA was minipreped from aliquots of cultures removed at desired time points. Restriction patterns of the DNAs from the primary cultures were compared with those from subsequent cultures. By digesting the DNAs with EcoRI and resolving the fragments in agarose gels, we could detect changes in 7 out of 18 Lawrist clones, and 7 out of 16 Supercos clones. Some examples of the instability are shown in Figure 2. Most of the changes in Lawrist 16 and Supercos clones occurred in 20–40 generations after primary culture, and appeared to be large deletions as judged by the disappearance of major bands present in 0 generations (see the clones 5, 11, and 17 of Lawrist, and clone 16 of Supercos). This dramatic reduction in cosmid size was not apparent in some other clones, such as the clones 4 and 14 (Figure 2C), in which the size of a few restriction fragments were changed probably due to rearrangements. By the same criterion, we found that 1 (Figure 2A, clone 8) out of 18 Fosmid clones underwent detectable change. This change does not appear to be a deletion but rather a rearrangement, resulting in the increase of the size of an EcoRI fragment. Further analysis of these Fosmid clones by HindIII digestion also detected change only in the clone 8 but not in the other 17 clones (not shown).

Cosmid clones containing mammalian DNA inserts can be highly unstable during routine maintenance and propagation, and frequently undergo drastic changes including large deletions that often generate low molecular weight DNA species. These can easily be detected by running the miniprep DNA in agarose gels. So far we have minipreped over a hundred human Fosmid clones and found no such deletions (not shown). From these results we conclude that Fosmid vector can propagate mammalian genomic DNA inserts with significantly greater stability than that of conventional multicopy cosmid vectors.

To directly show whether the inserts that are unstable in conventional vectors can be stably propagated in Fosmid vector, we cut out human inserts from the DNAs from the primary cultures of three unstable Supercos clones (0 generations of the clones 4, 14, and 16; Figure 2C) and recloned them into pFOS1. Fosmid colonies obtained as described in MATERIALS AND METHODS were picked and propagated to 40 and 80 generations, and analyzed by EcoRI and BamHI digestion. As shown in Figure 3, all three inserts recloned in the Fosmid vector underwent no detectable change during propagation.

In this report, we have demonstrated relative stability of human DNA inserts in Fosmid, a low copy number cosmid vector, during serial passage. The capability of Fosmid to stably propagate complex inserts has several major implications. First, successful use of genomic libraries, especially for detailed characterization and eventual sequencing of large genes of higher organisms, depends on the stability of the clones during propagation. Frequently, the characterization of clones occurs at a later stage and often in the laboratories other than the one in which the original clones were isolated. Hence, there will be ample opportunities for rearrangements during transfer of the clones between laboratories, although the undeleted clones can be rescued (1) provided the instability is recognized in time. Second, considerable segments of genomes from higher organisms have been known to be unclonable to *E.coli* mainly due to extreme instability in current high copy number vectors. It is conceivable that such sequences are clonable in Fosmid. Therefore, use of Fosmid vector that can maintain and propagate previously unstable or unclonable genomic segments should allow construction of libraries with fuller representation of genomes.

ACKNOWLEDGEMENTS

This work was supported by a grant from the Department of Energy, DE-FG03-89 ER60891 to M.I.S. Work by P.J.de Jong was supported by U.S. Department of Energy contract No.W-7405-ENG-48.

REFERENCES

1. Yokobata,K., Trenchak,B., and de Jong,P.J. (1991) Nucleic Acids Res. **19**,403–404.
2. Crowder,P.J., Cartwight,A.L., Hocking,A., Jefferson,S., Ford,M.D., and Woodcock,D.M. (1989) Nucleic Acids Res. **17**,7229–7239.
3. Nader,W.F., Edlind,T.D., Huettermann,A., and Sauer,H.W. (1985) Proc. Natl. Acad. Sci. U.S.A. **82**, 2698–2702.
4. Wyman,A.R. and Wertman,K.F. (1987) Methods Enzymol. **152**, 173–180.
5. Wertman,K.F., Wyman,A.R., and Botstein,D. (1986) Gene **49**, 253–256.
6. Sambrook,J., Fritsch,E.F., and Maniatis,T. (1989) Molecular Cloning, Cold Spring Harber.
7. Bernard,H.-U., and Helinski,D.R. (1979) Methods Enzymol. **68**,482–492.
8. Bates,P. (1987) Methods in Enzymol. **153**,82–94.