Gene, 106 (1991) 97–101

GENE 06066

Plasmids for recombination-based screening

(DNA sequencing; *Escherichia coli*; multiplexing; R6K, recombinant DNA; *supF*; suppressor)

Gordon D. Stewart, Michael A. Hauser, Harold Kang, Damien P. McCann, Małgorzata M. Osemlak, David M. Kurnit and Andrzej J. Hanzlik

Departments of Pediatrics and Human Genetics, Howard Hughes Medical Institute, University of Michigan Medical Center, Ann Arbor, MI 48109-0650 (U.S.A.)

Received by J.L. Slightom: 6 June 1991
Accepted: 1 July 1991
Received at publishers: 30 July 1991

SUMMARY

To facilitate recombination-based screening, we constructed the ColE1-based plasmid, πG4, that confers chloramphenicol resistance, contains a polylinker with multiple unique restriction enzyme recognition sequences, and contains the genetic marker, *supF*. To facilitate recombination-based screening followed by rapid DNA sequencing, we inserted the selectable marker, *supF*, into each of 20 high-copy-number (hcn) pUC-derived NoC plasmids that were designed for multiplex DNA sequencing. To facilitate recombination-based screening of common cDNA libraries that often contain ColE1 sequences, we constructed a *supF*-carrying plasmid whose replication was driven from an R6K replicon that does not share sequence homology with ColE1. Furthermore, we incorporated a useful polylinker and increased the copy number of this plasmid to create the 4.4-kb hcn plasmid, pMAD1. Thus, these plasmids allow: (*1*) background-free transformation of cells by a *supF* plasmid carrying an antibiotic-resistance marker; (*2*) simultaneous performance of the recombination-based assay and DNA sequencing; and (*3*) screening bacteriophage cDNA libraries that contain ColE1 sequences by recombination with a *supF* plasmid that is not homologous to ColE1 derivatives.

INTRODUCTION

Screening bacteriophage libraries by recombination is a rapid technique that promises to play a significant role in the deciphering of complex genomes (Seed, 1983; Kurnit and Seed, 1990). To facilitate this technique, we constructed three types of plasmids. First, we constructed a *supF* ColE1 plasmid vector that conferred Cm resistance directly. This plasmid, which has multiple unique cloning sites in the polylinker, avoids the use of an ancillary plasmid to select for *supF*. Second, the advent of multiplex sequencing processes establishes an efficient design for rapid DNA sequencing (Church and Kieffer-Higgins, 1988). We took advantage of this design to construct *supF*-containing plasmids that would satisfy both multiplex sequencing and recombination-based assay requirements. Third, screening bacteriophage libraries by recombination requires that the libraries are free of the vector sequences used for screening. Because of the presence of ColE1 sequences in λgt11 (Young and Davis, 1983), λgt10 (Huynh et al., 1985) and phage λ libraries constructed from YAC (Burke et al.,

Correspondence to: Dr. D.M. Kurnit, Howard Hughes Medical Institute, University of Michigan Medical Center, 1150 W. Medical Center Dr., Ann Arbor, MI 48109-0650 (U.S.A.);
Tel. (313)747-4747; Fax (313)936-9353.

Abbreviations: am, amber stop codon; Ap, ampicillin; bp, base pair(s); cDNA, DNA complementary to RNA; Cm, chloramphenicol; hcn, high copy number; IPTG, isopropyl-β-D-thiogalactopyranoside; kb, kilobase(s) or 1000 bp; Km, kanamycin; LB, Luria–Bertani (broth); nt, nucleotide(s); oligo, oligodeoxyribonucleotide; pfu, plaque-forming unit(s); ^{R}, resistance/resistant; *supF*, gene encoding mutant tyrosine tRNA that translates an am; Tc, tetracycline; wt, wild type; XGal, 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside; [ ], denotes plasmid-carrier state.

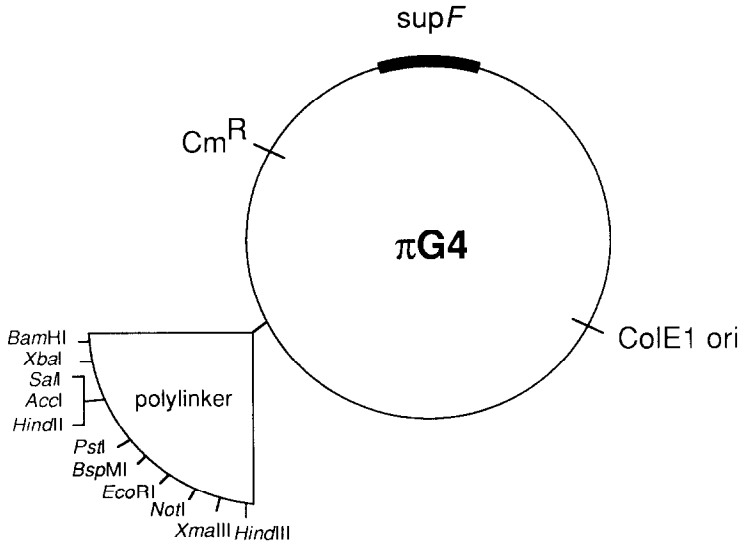
1987) recombinants, background-free recombination-based screening of these libraries requires the construction of a plasmid that carries *supF* but lacks sequence overlap with ColE1. We describe herein the elaboration of such a plasmid based on the R6K replicon.

EXPERIMENTAL AND DISCUSSION

(a) Creation of a *supF* plasmid that confers antibiotic resistance

To ensure the small size of *supF* plasmids, originally no antibiotic-resistance marker was placed directly on the plasmids; rather the *supF* marker was selected by requiring suppression of am mutations in antibiotic-resistance encoding genes on another plasmid, p3 (Km^R Ap^Ram Tc^Ram; Seed, 1983). While this was beneficial when very small *supF* plasmids were required, this strategy was plagued by the occurrence of bacterial chromosome suppressors. Although the frequency of occurrence of such chromosomal suppressors is low, the frequency of recombinant plasmid construction is often low enough or absent so that the cells that grow after selection for *supF* have the chromosomal, rather than the desired plasmid, suppressor tRNA.

To circumvent this difficulty, we cloned the Cm^R marker directly onto a plasmid that expresses *supF*. To accomplish this, a *Hae*II fragment expressing Cm^R was taken from pACYC184 (Chang and Cohen, 1978). This fragment was blunt-ended using mung bean nuclease and cloned into the blunt-ended *Sma*I site in the polylinker of πAN13 (Lutz et al., 1987), yielding πFLEE6. This vector has two *Eco*RI sites, one in the polylinker and one in the Cm^R gene. To eliminate the *Eco*RI site in the Cm^R gene of πFLEE6, we substituted the 339-bp *Dra*I fragment of this gene from pMLC28 (gift of B. Seed; Burmeister and Lehrach, 1988). Plasmid pMLC28 has this *Eco*RI site removed by directed mutagenesis of a single bp but still expresses Cm^R. Removal of the *Eco*RI site in the Cm^R gene of the resultant plasmid, πG13, was confirmed by restriction digestion and DNA sequencing (data not shown). A synthetic oligo was inserted into the *Hin*dIII site in the polylinker of πG13 to construct πG2-5. This oligo furnished *Not*I and *Eco*RI sites and recreated the unique *Hin*dIII site. The *Eco*RI site on the other side of the Cm^R element was removed by digesting with *Xmn*I + *Sac*I, blunt-ending the *Sac*I overhang with T4 DNA polymerase, and ligating the blunt-ended fragment.



Polylinker sequence:

SalI
HindII
AccI
XbaI
BspMI
NotI
BamHI
PstI
EcoRI
XmaIII
HindIII

5'-GGATCCTCTAGAGTCGACCTGCAGCCCAAGCTATCGAATTCTAGCGGCCGCTCGAAGCTT-3'

Fig. 1. The organization of πG4. This vector was constructed starting from πAN13 (Lutz et al., 1987). Plasmid πG4 has the following features: a ColE1 origin of replication, a Cm^R element, a *supF* gene and a polylinker. The sequence of the polylinker is given.

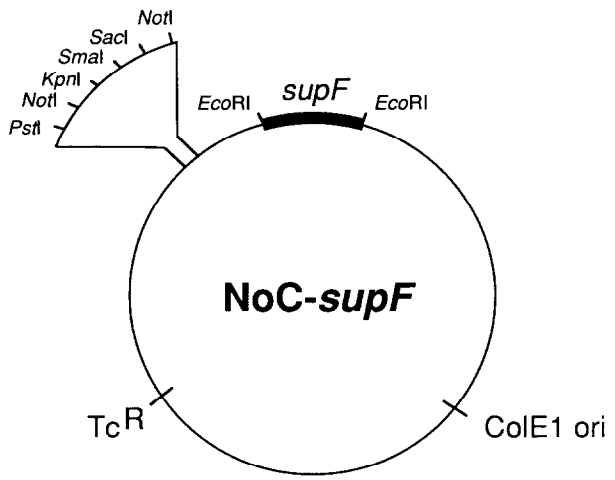


Fig. 2. The organization of NoC-*supF* vectors. These vectors were constructed starting from the NoC plasmids described by Church and Kieffer-Higgins (1988). In each case, the plasmid was opened at the unique *Eco*RI site and a 203-bp *Eco*RI fragment from pGFYπVX (section **b**) spanning the *supF* gene was inserted.

This engendered the loss of 10 bp including the concomitant destruction of *Eco*RI, *Xmn*I and *Sac*I recognition sites. Sequencing was performed to ensure that all the changes made to construct πG4 were done correctly. The unique cloning sites in the polylinker are for *Acc*I, *Bam*HI, *Bsp*MI, *Eco*RI, *Hin*dII, *Hin*dIII, *Not*I, *Pst*I, *Sal*I, *Xba*I and *Xma*III (Fig. 1).

(b) Insertion of *supF* into plasmids designed for multiplex sequencing

Rapid multiplex sequencing demanded the creation of hcn plasmids with cloning sites surrounded by different oligos that allowed sequential multiplex sequencing. To meet this need, Church and Kieffer-Higgins (1988) elaborated the NoC family of plasmids which abuts the cloning site of pUC-based hcn plasmids with synthetic oligos that permit primed sequencing analyses. We inserted *supF* into the *Eco*RI site of each of the 20 distinct NoC plasmids (Fig. 2). The *supF* gene was derived from pGFYπVX, constructed by inserting a *Pst*I fragment encompassing πVX (Seed, 1983) into the *Pst*I site of pGFY218 (Amann et al., 1983). This construct grew better than the original πVX plasmid and was therefore used as the source for *supF*. An *Eco*RI digest of pGFYπVX yielded a 203-bp fragment spanning *supF* that was ligated into the *Eco*RI site of each of the NoC plasmids (Church and Kieffer-Higgins, 1988). Twenty ligations were done separately to clone *supF* into each of the 20 distinct NoC plasmids. The ligations were transformed into MC1061[p3], containing the Km^R Ap^Ram Tc^Ram plasmid, p3 (Seed, 1983). The presence of *supF* was then ascertained by a Km^R Ap^R Tc^R phenotype (Seed, 1983). In addition, each cloning product was checked for functional *supF* by the ability of the product to support plating of λ phages carrying am mutations in the A and B genes (Kurnit and Seed, 1990). The result is a series of NoC-*supF* plasmids that permit both multiplex sequencing and recombination-based screening. The viability of these strains demonstrates that presence of a hcn of *supF* is tolerated.

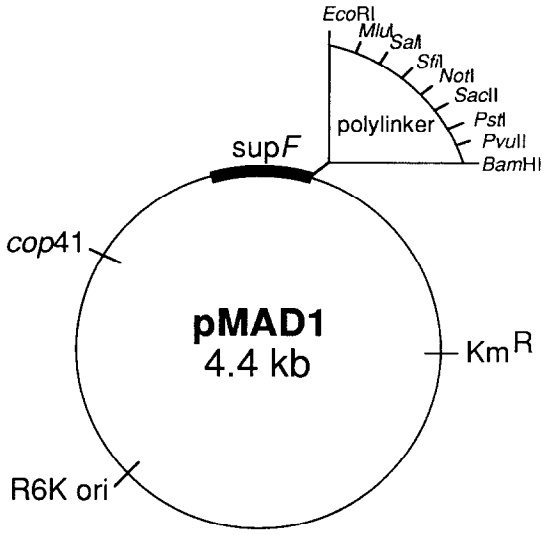
(c) Insertion of *supF* into an R6K replicon

For recombination-based screening (Seed, 1983; Kurnit and Seed, 1990), there can be no shared nt sequence identity between the screening plasmid and the library to be screened which would result in false positives. Unfortunately, λgt11 libraries (Young and Davis, 1983) require the presence of the ColE1 plasmid, pMC9, to suppress expression of the *lac* promoter in λgt11. Recombination between phages and this plasmid mediated by shared *lac* sequences (at a frequency of approx. 10^{-4}) results in the integration of ColE1 (pMC9) sequences into the phages. This prevents background-free recombination-based screening of these libraries with plasmids that contain the ColE1 origin. Although λgt10 libraries should not suffer this problem theoretically, we have found that many λgt10 libraries do contain ColE1-derived sequences (Jankowski et al., 1990), indicating that these libraries have been passaged on the strains intended for λgt11 or that accidental contamination has occurred. Regardless, ColE1-derived replicons cannot be used for recombination-based screening of these λgt10 cDNA libraries. Vectors carrying YAC DNA contain ColE1 sequences that are present in the YAC vehicle (Burke et al., 1987). This prevents background-free recombination-based screening of libraries constructed from YAC vehicles with a sequence inserted in a ColE1 origin plasmid.

TABLE I

Sequences of oligos used herein to construct pMAD1

5′-GTTTCGGACTTTTGAAAG-3′; (end of *supF* without polylinker attached); (18 bp)
5′-AATTGTTTCGGACTTTTGAA-3′; mutation underlined; (end of *supF* without polylinker attached); (20 bp)
5′-AATTGGATCCAGCTGCAGCGGCCGCGGAGGCCGTCGACGCGTGAATTCTTTCTCAACGTAACACTTTAC-3′; mutation underlined; (end of *supF* with polylinker attached); (69 bp)
5′-GGATCCAGCTGCAGCGGCCGCGGAGGCCGTCGACGCGTGAATTCTTTCTCAACGTAACACTTTAC-3′; (end of *supF* with polylinker attached); (65 bp)



SfiI
MluI NotI PvuII
EcoRI SalI SacII PstI BamHI

GTGTTACGTTGAGAAAGAATTCACGCGTCGACGGCCTCCGCGGCCGCTGCAGCTGGATCCAATTCCGGAT

3' end of supF polylinker

Fig. 3. The organization of pMAD1. This plasmid was constructed starting from the R6K-derived element, pcos2EMBL (Poustka et al., 1984). The large *Eco*RI fragment of pcos2EMBL was ligated to a PCR segment that spanned the *supF* gene, destroyed these *Eco*RI sites and generated a polylinker whose sequence is given. This intermediate was termed pAD1. The small *Hgi*AI-*Ban*I fragment of pAD1 was replaced by the small *Hgi*AI-*Ban*I fragment of pGN7-*cop*41 (provided by M. Inuzuka and P. Wada; Inuzuka and Wada, 1985) to create the hcn plasmid, pMAD1. The overall structure and polylinker of this plasmid are given.

To eliminate this background, we cloned *supF* into an R6K plasmid replicon which is not homologous to ColE1 (Kolter and Helinski, 1978; Poustka et al., 1984). Again, the *Eco*RI fragment encompassing the *supF* gene in pGFYπVX was the source of *supF* for the construction of this R6K-derived vector. Two PCR amplifications (Saiki et al., 1985) were performed using the oligo primers in Table I to construct equivalent length fragments that spanned the *supF* gene of pGFYπVX and that were offset by 4 bp. In addition to amplifying the *supF* gene, the oligo primers used for these PCR reactions created a polylinker containing sites for *Bam*HI, *Eco*RI, *Mlu*I, *Not*I, *Pst*I, *Pvu*II, *Sac*II, *Sal*I and *Sfi*I (Fig. 3). When the products of these two PCR amplifications were denatured and annealed (Jones et al., 1990) the resultant mixture contained both blunt-ended and staggered-ended molecules. By design, the latter had 5′–AATT–3′ overhangs compatible with *Eco*RI overhangs. Further, the oligos used to prime the PCR (Table I) were constructed so that a 1-bp mismatch resulted in the destruction of intact *Eco*RI sites at each end by mutation of the 3′-most nt of the *Eco*RI recognition sequence from C to G. However, the continued presence of 5′–AATT–3′ ensured that each overhang would still ligate with an *Eco*RI overhang. These *supF*-polylinker fragments generated by PCR lacked 5′-phosphates by virtue of their mode of synthesis. Thus, a large excess of these fragments could be added without ligation occurring among them. An *Eco*RI digest of pcos2EMBL (Poustka et al., 1984) resulted in the liberation of the R6K replicon and a Km^R element together lacking identity to pBR322 or phage λ (Poustka et al., 1984). This *Eco*RI fragment was ligated to the sticky-ended *supF* fragment generated during the two PCRs followed by denaturation and annealing. After transformation into the Su^0 *lacZ*am host, LG75 (Guarente et al., 1980), Km^R Tc^S colonies that were blue in the presence of IPTG and XGal had the desired Km^R Tc^S *supF* genotype with the designed polylinker. The plasmid in one such colony was termed pAD1.

Increasing the copy number of pAD1 required insertion of a *cop*41 mutation (Inuzuka and Wada, 1985). This single

bp change was shown previously to increase the copy number of R6K from approx. 20 to approx. 100 (Inuzuka and Wada, 1985). The *cop*41 mutation was isolated on a *Ban*I-*Hgi*AI restriction fragment from pGN7 *cop*41 (Inuzuka and Wada, 1985). This fragment was ligated to the large *Ban*I-*Hgi*AI fragment of pAD1. Following transformation into the *lacZ*am Su^0 host, LG75, the Km^R *supF* plasmid, pMAD1, was isolated with this hcn *cop*41 mutation.

To demonstrate that the *cop*41 mutation was present, we distinguished the wt sequence from the *cop*41 sequence by the presence of an *Rsa*I site engendered by the normal sequence but not by the mutant. The region in question was amplified by PCR using two oligos, one of which abutted the *cop*41 mutation and the other was in *supF*. *Rsa*I digestion of this PCR fragment yielded a 220-bp fragment if the wt sequence was present and a 240-bp fragment if the mutant *cop*41 sequence was present. This analysis confirmed that pMAD1 contained the desired *cop*41 mutation.

To underscore that pMAD1 lacks identity to ColE1 sequences, we showed by recombination (Seed, 1983; Kurnit and Seed, 1990) that this plasmid does not recombine with ColE1 nt sequences in λ phages. This lack of identity is consistent with the known sequence of R6K and with the finding that R6K-based replicons do not recombine with cosmids carrying ColE1 sequences (Poustka et al., 1984). In summary, pMAD1 is a 4.4-kb hcn R6K-derived plasmid that confers Km^R and *supF*, has a polylinker with multiple cloning sites and is not homologous to ColE1 plasmids (Fig. 3). This plasmid can be used to screen libraries fabricated in phage λ regardless of whether the ColE1 replicon is present therein. To illustrate this point, we have used this plasmid to study the transcription pattern of more than 500 genomic sequences (M.A.H., M.M.O., D.M.K. and A.J.H., unpublished).

The result of these constructions is that the applicability of the recombination-based assay is broadened. In particular, the last application permits us to screen a wide variety of cDNA libraries constructed in vectors that have ColE1 sequences. This increases the number and kind of libraries that are amenable to the recombination-based approach.

ACKNOWLEDGEMENTS

We thank Brian Seed for providing pMLC28, Hans Lehrach for providing pcos2EMBL, and George Church for providing the NoC plasmids. Frederick Lee constructed the πFLEE6 plasmid. M.A.H., M.M.O. and A.J.H. are Associates and D.M.K. is an Investigator of the Howard Hughes Medical Institute. Supported also by a grant from the Alzheimer's Disease and Related Disorders Association to G.D.S.

REFERENCES

Amann, E., Brosius, J. and Ptashne, M.: Vectors bearing a hybrid *trp-lac* promoter useful for regulated expression of cloned genes in *Escherichia coli*. Gene 25 (1983) 167–178.

Burke, D.T., Carle, G.F. and Olson, M.V.: Cloning of large segments of exogenous DNA into yeast by means of artificial chromosome vectors. Science 236 (1987) 806–812.

Burmeister, M. and Lehrach, H.: Jekyll, a family of phage-plasmid shuttle vectors. Gene 73 (1988) 245–250.

Chang, A.C.Y. and Cohen, S.N.: Construction and characterization of amplifiable multicopy DNA cloning vehicles derived from the P15A cryptic miniplasmid. J. Bacteriol. 134 (1978) 1141–1156.

Church, G.M. and Kieffer-Higgins, S.: Multiplex DNA sequencing. Science 240 (1988) 185–188.

Guarente, L., Lauer, G., Roberts, T.M. and Ptashne, M.: Improved methods for maximizing expression of a cloned gene: a bacterium that synthesizes rabbit β-globin. Cell 20 (1980) 543–553.

Inuzuka, M. and Wada, Y.: A single amino acid mutation in the initiation protein is responsible for the DNA overproduction phenotype of copy number mutants of plasmid R6K. EMBO J. 4 (1985) 2301–2307.

Jankowski, S., Stewart, G.D., Buraczynska, M., Galt, J., Van Keuren, M. and Kurnit, D.M.: Molecular approaches to trisomy 21. In: Epstein, C. and Patterson, D. (Eds.), 21st Chromosome and Down Syndrome, Wiley-Liss, New York, 1990, pp. 79–88.

Jones, D.H., Sakamoto, K., Vorce, R.L. and Howard, B.H.: DNA mutagenesis and recombination. Nature 344 (1990) 793–794.

Kolter, R. and Helinski, D.R.: Construction of plasmid R6K derivatives in vitro: characterization of the R6K replication region. Plasmid 1 (1978) 571–580.

Kurnit, D.M. and Seed, B.: Improved genetic selection for screening bacteriophage libraries by homologous recombination in vivo. Proc. Natl. Acad. Sci. USA 87 (1990) 3166–3169.

Lutz, C.T., Hollifield, W.C., Seed, B., Davie, J.M. and Huang, H.V.: Syrinx 2A: an improved λ phage vector designed for screening DNA libraries by recombination in vivo. Proc. Natl. Acad. Sci. USA 84 (1987) 4379–4383.

Poustka, A., Rackwitz, H.-R., Frischauf, A., Hohn, B. and Lehrach, H.: Selective isolation of cosmid clones by homologous recombination in *Escherichia coli*. Proc. Natl. Acad. Sci. USA 81 (1984) 4129–4133.

Saiki, R.K., Scharf, S., Faloona, F., Mullis, K.B., Horn, G., Erlich, H.A. and Arnheim, N.: Enzymatic amplification of β-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. Science 230 (1985) 1350–1354.

Seed, B.: Purification of genomic sequences from bacteriophage libraries by recombination and selection in vivo. Nucleic Acids Res. 11 (1983) 2427–2445.

Young, R.A. and Davis, R.W.: Efficient isolation of genes by using antibody probes. Proc. Natl. Acad. Sci. USA 80 (1983) 1194–1198.