

pFOS-LA: a modified vector for production of random shear fosmid libraries

Jonathan L. Longmire and Nancy C. Brown
Los Alamos National Laboratory, Los Alamos, NM, USA

BioTechniques 35:50-54 (July 2003)

Fosmid vectors have proven to be valuable in genomic studies because they contain single-copy origins of replication and allow stable propagation of DNA inserts in the size range of 30–40 kb (1–5). The original fosmid vector, pFOS-1, allowed high-efficiency cloning of target DNA fragments produced by partial digestion (6). pFOS-1 has the advantages of being a double cos site vector, which include (i) cloning arms can be prepared easily and (ii) target fragments can be dephosphorylated and therefore do not need to be size-selected before cloning. The cloning region of pFOS-1 contains *Bam*HI and *Hind*III sites that partially digested and dephosphorylated target fragments can be cloned into. *Bam*HI is perhaps the most useful of these two sites because of the large number of 4-base cutting restriction enzymes that leave ends that are compatible with the cohesive end left by *Bam*HI (such as *Mbo*I and *Sau*3AI).

A distinct limitation of the original pFOS-1 vector was the lack of unique blunt end sites in the cloning region that would allow cloning of DNA fragments generated by random shearing. Such libraries have potential advantages over partial digest libraries in terms of sequence representation. This is especially true in cases where the target genomes may not have random distributions of restriction sites and/or where these sites are protected by methylation (as is the case for many microbial genomes). Here we describe modifications made to pFOS-1 to allow the production of random shear fosmid libraries. The modified vector is termed pFOS-LA.

The following oligonucleotide sequences were designed to provide unique blunt end sites in the cloning region: 5'-AGCTGGATTTAAATCACGTG-3' and 5'-AGCTCACGTGATTTA-

AATCC-3'. When double-stranded, these oligonucleotides have termini that are cohesive with the ends of *Hind*III-linearized pFOS-1. However, the *Hind*III site is not regenerated upon insertion of the oligonucleotide duplex into the vector. The central region of the duplex contains recognition sites for two restriction enzymes that leave blunt ends, *Swa*I and *Pml*I.

The modified vector pFOS-LA was derived as follows (Figure 1). Intact pFOS-1 DNA was linearized at the *Hind*III site and ligated to a 10-fold molar excess of pre-annealed, complementary oligonucleotides described above. Ligation products were transformed into *Escherichia coli* host strain

DH5 α MCR and plated on LB agar containing 25 μ g/mL chloramphenicol. Several individual colonies were picked for plasmid DNA miniprepping, and the desired recombinants were identified through test cuts that demonstrated (i) the lack of a *Hind*III site and (ii) the presence of unique recognition sites for both *Swa*I and *Pml*I. The *Bam*HI site within the cloning region of pFOS-1 was preserved in the modified vector to maintain the capability for production of partial digest libraries (if so desired).

To prepare cloning arms, circular pFOS-LA DNA was digested with *Aat*I and treated with calf intestinal phosphatase, followed by phenol:chloroform extraction and microdialysis against TE buffer (10 mM Tris-HCl, pH 8.0, 1 mM EDTA). Cloning arms were made by digesting the linearized and dephosphorylated vector DNA with *Swa*I, followed by another round of phenol:chloroform extraction and microdialysis. Aliquots (200 ng) of the fully prepared, blunt-end cloning arms were stored frozen at -20°C.

A random shear, total genomic library was made in pFOS-LA for the mi-

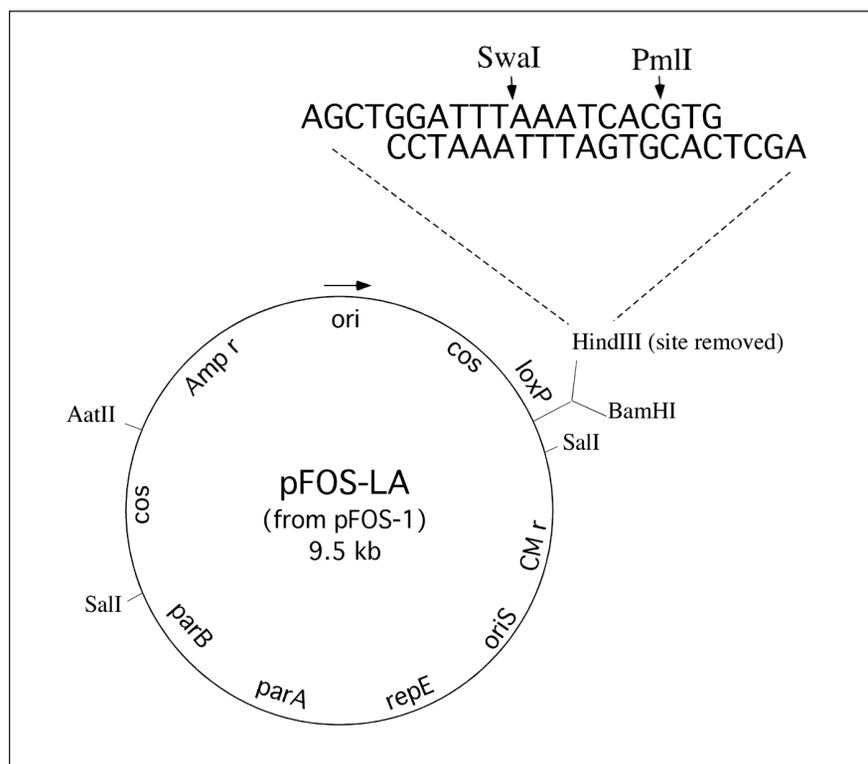


Figure 1. Derivation of pFOS-LA from pFOS-1 (adapted from Reference 6).

Table 1. Characteristics of Random Shear *N. punctiforme* Fosmid Library

Vector	DNA Fragmentation	DNA Mass Cloned	Total cfu	Cloning Efficiency ^a	\bar{x} Insert Size ^b	Genome Rep. ^c
pFOS-LA	Random shear	200 ng	95,000	4.8×10^5	41.3 kb	413×

^acfu/ μ g input DNA.
^bBased on *Eco*RI fingerprinting and fragment sizing 96 randomly selected clones.
^cBased on average insert size and an estimated *N. punctiforme* genome size of 9.5 Mbp.

crobe *Nostoc punctiforme*. Solution-purified, genomic DNA from *N. punctiforme* (2 μ g in 100 μ L TE buffer) was pushed through a 13-mm, 0.2- μ m filter (NALGENE[®]; Nalge Nunc International, Rochester, NY, USA) using a 1-cc syringe. This filtering served to shear the DNA to a homogeneous size that was appropriate for fosmid cloning (approximately 40 kb). The sheared DNA was end-repaired to make blunt ends (End-It[™] kit; EPICENTRE, Madison, WI, USA), followed by phenol:chloroform extraction and microdialysis against TE. The DNA was subsequently dephosphorylated using calf intestinal

phosphatase, followed by cleanup with phenol:chloroform extraction and microdialysis. For actual cloning, 200 ng of the processed *N. punctiforme* target DNA was ligated overnight at 16°C in the presence of 4.5% polyethylene glycol to 400 ng pFOS-LA cloning arms prepared with *Swa*I blunt cloning ends. Ligation products were extracted using phenol:chloroform, microdialysed, and packaged in vitro using Gigapack[®] Gold extracts (Stratagene, La Jolla, CA, USA). Packaging extracts were used to infect *E. coli* host strain DH5 α MCR. Infected cells were plated on LB agar containing 25 μ g/mL chloramphenicol.

As summarized in Table 1, the resulting library contained approximately 95,000 clones for a cloning efficiency of 4.8×10^5 cfu/ μ g input DNA. Given a measured average insert size of approximately 41.3 kb, the library was roughly 413-fold representative for the 9.5-Mbp genome of this microbe (7). Restriction fingerprints of 96 randomly selected clones revealed a high level of diversity among the inserts and no apparent preferential cloning (Figure 2).

A portion (10-fold clone coverage) of the *N. punctiforme* pFOS-LA library was end-sequenced as part of a total genome sequencing effort for this microbe (7). Analyses of those data demonstrate that the pFOS-LA clones fall evenly onto the assembly generated from shot-

gun sequencing total genomic plasmid libraries (S. Malfatti, Lawrence Livermore National Laboratory, personal communication). This observation provides additional evidence that the pFOS-LA library is highly representative for the *N. punctiforme* genome and indicates the significant utility of this library in the final sequencing of *N. punctiforme* (to be the subject of subsequent papers).

In addition to producing libraries with possibly better sequence representation, there are certain other advantages in making random shear fosmid libraries. First, the requirement for isolating very high molecular weight target DNA is less stringent. When making fosmid (or cosmid) libraries by partial digestion, the starting DNA has to be very large (>100 kb). This is due to the requirement for the partial digestion reactions to generate 30- to 40-kb fragments with cohesive sites at both ends. Because of this size requirement, the efficiency of making partial digest libraries drops dramatically with smaller starting DNA (8). The rule of “bigger is better” definitely applies in the case of cloning by partial digestion.

For making random shear libraries, on the other hand, the starting DNA can be much smaller and only needs to be the same size as the final insert fragments at 30–40 kb. Obtaining fragments of this size is quite simple because most protocols that are used for isolating genomic DNA in solution typically yield DNA fragments in this size range. Another advantage of making random shear libraries is they are much simpler to build, since partial digestion reactions (which can be tedious and time-consuming) are not needed. Of the two approaches, the production of random shear fosmid libraries is much more amenable to high throughput and possible automation.

Cloning random shear fragments into pFOS-LA required blunt-end ligation (insert to vector), which is not nearly as efficient as ligation of cohesive ends as used in making partial

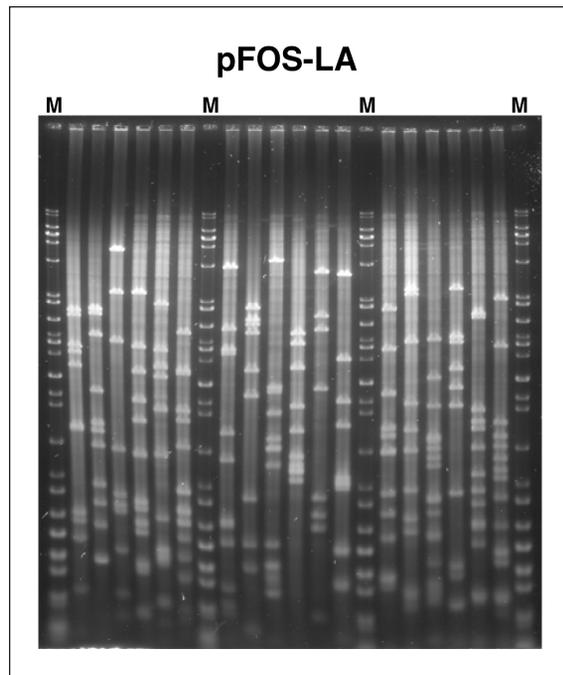


Figure 2. Diversity of inserts in the *N. punctiforme* fosmid library made by random shear. Randomly selected clones (96 analyzed, 18 are shown) were digested with *Eco*RI before agarose gel electrophoresis and fragment sizing. Lanes marked M contain size standards comprising a mixture of uncut and *Hind*III-digested λ DNA, *Hae*III-digested θ X 174 DNA, and 1-kb DNA ladder.

digest libraries. Therefore, a possible concern is that the cloning efficiency in the random shear approach may be too low to be practical and that large amounts of target DNA would be needed to produce multiple representation fosmid libraries. However, results obtained in the construction of the *N. punctiforme* library clearly demonstrate that 30- to 40-kb fragments produced by random shearing (followed by end repair and dephosphorylation) can be cloned into pFOS-LA at high efficiency ($>10^5$ cfu/ μ g target DNA). This level of cloning efficiency is clearly adequate for producing genomic libraries of high statistical representation from microgram and submicrogram quantities of DNA. The cloning efficiency observed in these experiments, taken together with a tolerance for smaller (even slightly degraded) starting DNA, enables and encourages the use of pFOS-LA for making libraries from small amounts of DNA obtained from animal and plant tissues, microbial material, and flow-sorted chromosomes.

ACKNOWLEDGMENTS

The authors thank Dr. Jack Meeks (University of California at Davis), Patrick Chain, and Stephanie Malfatti (Lawrence Livermore National Laboratory) for providing genomic DNA and information regarding the sequence assembly of *N. punctiforme*. This work was conducted under the auspices of the USDOE. The modified vector pFOS-LA is available upon request.

REFERENCES

1. Dauwerse, J.G., J.W. Wessels, R.H. Giles, J. Wiegman, B.A. Vanderreiden, G. Fugazza, E.A. Jumelet, E. Smit, et al. 1993. Cloning the breakpoint cluster of the *inv(16)* in acute nonlymphocytic leukemia M4 EO. *Hum. Mol. Genet.* 2:1527-1534.
2. Gingrich, J.C., D.M. Boehrer J.A. Garnes, W. Johnson, B.S. Wong, A. Bergmann, G.G. Eveleth, R.G. Langlois, and A.V. Carrano. 1996. Construction and characterization of human chromosome 2-specific cosmid, fosmid, and PAC clone libraries. *Genomics* 32:65-74.
3. Chibana, H., B.B. Magee, S. Grindle, Y. Ran, S. Scherer, and P.T. Magee. 1998. A physical map of chromosome 7 of *Candida albicans*. *Genetics* 149:1739-1752.
4. Kingzette, M., H. SpiekerPolet, P.C. Yam, S.K. Zhai, and K.L. Knight. 1998. Trans-chromosomal recombination within the Ig heavy chain switch region in B lymphocytes. *Proc. Natl. Acad. Sci. USA* 95:11840-11845.
5. Quaiser, A., T. Ocsenreiter, H.P. Klenk, A. Kletzin, A.H. Treusch, G. Meurer, J. Eck, C.W. Sensen, and C. Schleper. 2002. First insight into the genome of an uncultivated crenarchaeote from soil. *Environ. Microbiol.* 4:603-611.
6. Kim, U.J., H. Shizuya, P.J. DeJong, B. Birren, and M.I. Simon. 1992. Stable propagation of cosmid sized human DNA inserts in an f-factor based vector. *Nucleic Acids Res.* 20:1083-1085.
7. Meeks, J.C., J. Elhai, T. Thiel, M. Potts, F. Larimer, J. Lamerdin, P. Predki, and R. Atlas. 2001. An overview of the genome of *Nostoc punctiforme*, a multicellular, symbiotic cyanobacterium. *Photosynthesis Res.* 70:85-106.
8. Longmire, J.L., N.C. Brown, L.M. Meincke, M.L. Campbell, K.L. Albright, J.F. Fawcett, E.W. Campbell, R.K. Moyzis, et al. 1993. Construction and characterization of partial digest DNA libraries made from flow-sorted human chromosome 16. *Genet. Anal. Tech. Appl.* 10:69-76.

Received 20 March 2003; accepted 17 April 2003.

Address correspondence to Jonathan L. Longmire, Mail Stop M888, Los Alamos National Laboratory, Los Alamos, NM, USA. e-mail: longmire@telomere.lanl.gov

ELAM-1/E-selectin promoter contains an inducible AP-1/CREB site and is not NF- κ B-specific

Liselotte E. Jensen and Alexander S. Whitehead

University of Pennsylvania School of Medicine, Philadelphia, PA, USA

BioTechniques 35:54-58 (July 2003)

The innate immune system plays an important role as a first defense against pathogens and involves the recognition of bacteria and viruses, and byproducts thereof, by Toll receptors on immune-competent cells (1). Activated cells synthesize and secrete cytokines, which in turn activate systemic responses directed at clearing the pathogen. Two important cytokines are interleukin-1 (IL-1) and tumor necrosis factor- α (TNF- α). These cytokines, as well as the Toll receptors, initiate intracellular signaling cascades that activate nuclear factor κ B (NF- κ B), activator protein-1 (AP-1), and cAMP responsive element binding proteins (CREBs), which are transcription factors essential for the regulation of numerous genes, many of which play important roles in immunological processes. Subunits of the AP-1 and the CREB protein families may form both homo- and heterodimers; the latter may comprise subunits both from within and between protein families. The signaling pathways leading to acti-

vation of at least NF- κ B and AP-1 are common proximal to the receptors but diverge downstream to provide specificity (2). A common tool for examining signaling cascades is the luciferase (or chloramphenicol acetyltransferase) reporter assay. Such assays offer advantages such as simple protocols, no requirement for highly efficient transfections, and the possibility of elucidating molecular orders of signaling factors and determining branch-points of diverging signaling cascades.

The endothelial leukocyte-adhesion molecule (ELAM)-1 (also called E-selectin) promoter contains three NF- κ B sites, two of which are partially overlapping, that have been reported to be required for full induction by cytokines. A putative AP-1 site at position -499 to -493 within the promoter has been shown not to affect cytokine induction (3). Consequently, this promoter is often used to drive expression of luciferase in reporter assays and is frequently considered to be an "NF- κ B-specific" promot-