Benchmarks

Vector for High-Throughput Sequencing: Construction and Preparation with Cyclic Cut-Ligation

BioTechniques 30:1208-1210 (June 2001)

At present, shotgun sequencing is the most widely used method for genomic high-throughput sequencing. A critical step in this approach is the subcloning of the large insert DNA. For large-scale genomic sequencing to be efficient, subclones in libraries must cover the target DNA at the molecular level evenly. Moreover, these libraries must contain a minimum number of clones that have no inserts or that have chimeric inserts.

The double-adapter method (1) has been shown to be the most appropriate for this effort. In the method, physically sheared and end-repaired fragments are ligated to oligonucleotide adapters creating 12-base overhangs. The vector is prepared from a modified M13 vector, by *KpnI/PstI* digestion, followed by ligation to oligonucleotides with ends complementary to the overhangs created in the digest. The vector adapters create 5' overhangs complementary to those on the inserts. Following annealing of inserts to vectors, the DNA is directly used for transformation without ligation.

To generate subclone libraries, it is possible to use either M13 phage vector or plasmid vector. Unlike the M13 template, however, the plasmid template can be sequenced from both ends. This characteristic facilitates the subsequent ordering and orientation of the sequence contigs and reduces the possibility of sequence assembling errors. Because of this advantage, we choose plasmid as the sequencing template.

We have previously used physical sharing and blunt-end ligation to subclone bacterial artificial chromosome (BAC) or cosmid DNAs. To take advantage of the double-adapter method, we have modified the phagemid pBluescript® SK- vector (Stratagene, La Jolla, CA, USA) (3) to make it compatible with this procedure.

To construct the double-adapter phagemid vector, we started with a pBluescript SK vector containing an

approximately 3-kb human chromosome 16 insert cloned at the *Bam*HI site. Part of the human genomic fragment would be used as a spacer in the new vector. The DNA of this subclone was prepared using a QIAGEN® Plasmid Mini Kit (Qiagen, Valencia, CA, USA).

The 3-kb human genomic insert and the vector fragment were very close in size and difficult to separate from each other by agarose gel electrophoresis. To reduce the size of this human insert, we evaluated the restriction digestion of the clone with a variety of enzymes that are present in the polylinker site. Eight aliquots of 2 µg DNA of this clone were digested with *KpnI*, *SacI*, *EcoRI*, *EcoRV*, *HindIII*, *XhoI*, *PstI*, and *SmaI* (New England Biolabs, Beverly, MA, USA) in the appropriate 1× buffer for 2 h in a total volume of 20 µL. A 1% low melting gel was used to separate the expected fragment from the insert. A *SmaI* site was found in the human genomic insert.

Having been separated with the *SmaI* fragment (with a length of 0.9

kb), the vector fragment was approximately 4.8 kb. A slice of the 4.8-kb fragment was melted at 65°C; 10 µL of this slice were added with 10 µL 2× ligation buffer and then ligated with itself with 100 U ligase (New England Biolabs) at 4°C overnight. DNAs from two transformants were prepared and digested with *SmaI*, *KpnI*, and *SacI*. With *SmaI*, *KpnI*, or *SacI* digestion individually, a single 4.8-kb fragment was obtained, respectively. When digested with *KpnI* and *SacI* in a single reaction, two fragments were obtained, one 2.0 kb and one 2.8 kb in size. The larger fragment was pBluescript SK with a deletion of the most of the polylinker. All of the restriction sites and SK primer were deleted. The primers left for sequencing were T7, T3, and M13-20 and reverse primers. Figure 1 shows a map of the modified pBluescript vector (pSV-20) before preparation.

To prepare this vector for double-adapter subcloning, we used a QIAGEN Plasmid Mega Kit on the DNA of pSV-20. Approximately 100 µg pSV-

Figure 1. Map of the modified pBluescript vector pSV-20. It was constructed by inserting an approximately 2-kb human chromosome 16 fragment into the *Bam*HI and *SmaI* restriction site of pBluescript SK.

Benchmarks

20 were digested with 500 U *KpnI* and 500 U *SacI* at 37°C for 30 min in 700 µL 1× ligase buffer. This was followed by adding 300 µL 1× ligase buffer with 4000 U ligase and 50 µg each of the two vector adapters, VA-Sac (5'-TGATGAACTACTAGCT-3') and VA-Kpn (5'-TGATGAACTACTGTAC-3'). After mixing, the reaction was divided into 10 PCR tubes and placed into a model 9600 thermal cycler (Applied Biosystems, Foster City, CA, USA). Cyclic cut-ligation reactions were carried out for 600 cycles of 37°C for 1 min and 16°C for 1 min. The reactions were run in a 1% agarose gel. The 2.8-kb vector fragment was isolated and purified with a QIAquick™ Gel Extraction Kit (Qiagen).

We first tested the vector by having it ligated with itself. Approximately 2 µg isolated vector DNA were ligated overnight with 10 U ligase at 4°C in a volume of 20 µL. No ligation was detected (Figure 2).

Second, the vector was tested in transformations. The inserts were prepared following the protocol described by Andersson et al. (1). In the procedure, DNA from a cosmid was sheared then end-repaired with T4 DNA polymerase and Klenow. Once the insert adapter, 5'-AGTAGTTCATCA-3', was annealed to helper adapter 5'-TGATGA-3', the sheared and end-repaired cosmid DNA fragments were ligated with the insert adapter.

Transformation was performed as described by Sambrook et al. (2). On average, transformations yielded approximately 2200 transformant/µg starting DNA. Ninety-six of the subclones were amplified with the primers that flanked the cloning site. All the subclones used for the test contained inserts with the expected size (0.8–1.2 kb). In the past year at our center, high-throughput sequencing of 55 BACs and 21 cosmids confirmed that more than 99.9% of the subclones obtained with the procedure contained inserts.

The modified phagemid vector pSV-20 takes advantage of the double-

adapter method, which originally was developed for the M13 phage vector. This method is well suited for high-throughput sequencing and offers the advantage of paired-end plasmid reads. The constructed libraries are evenly represented and readily yield sufficient clones to complete large-scale projects. The frequency of clones containing an insert is close to 100%. The singular step of cyclic cut-ligation vector preparation is a much easier and more efficient procedure. All of these elements make this approach very easy to automate, which in turn increases accuracy and efficiency even further. The efficiency of high-throughput sequencing with pSV-20 is compatible with that of sequencing with M13 vector while maintaining the advantage of paired-end plasmid sequencing.

REFERENCES

1. Andersson, B., M.A. Wentland, J.Y. Ricafrente, W. Liu, and R.A. Gibbs. 1996. A "double adapter" method for improved shotgun library construction. *Anal. Biochem.* 236:107-113.
2. Sambrook, J., E.F. Fritsch, and T. Maniatis. 1989. *Molecular Cloning: Plasmid Vectors*. CSH Laboratory Press, Cold Spring Harbor, NY.
3. Short, J.M., J.M. Fernandez, J.A. Sorge, and W.D. Huse. 1988. Lambda ZAP: a bacteriophage lambda expression vector with in vivo excision properties. *Nucleic Acids Res.* 16:7583-7600.

This project was supported by the US Department of Energy, OBER, under contract no. W-7405-ENG-36. Address correspondence to Dr. Cliff S. Han, Bioscience Division and Center for Human Genome Studies, Mail Stop M888, Los Alamos National Laboratory, Los Alamos, NM 87545, USA. e-mail: chan@telomere.lanl.gov

Received 19 September 2000; accepted 26 March 2001.

Cliff S. Han, Judy Buckingham, Linda J. Meincke, and Norman A. Doggett
*DOE Joint Genome Institute
Los Alamos National
Laboratory
Los Alamos, NM, USA*

Figure 2. Agarose gel electrophoresis of the modified vector DNA prepared with cyclic cut-ligation. M, λ DNA/*HindIII*; lane 1, pSV-20 cut with *KpnI* and *SacI*; lane 2, pSV-20 prepared with cyclic cut-ligation reaction (see text for details); lane 3, the isolated 2.8-kb vector fragment; lane 4, no ligation was detected between the isolated vector fragments prepared with cyclic cut-ligation.