Short Technical Reports

Rapid Construction of Sequencing Templates by Random Insertion of Antibiotic Resistance Genes

BioTechniques 32:300-304 (February 2002)

ABSTRACT

We describe a novel and handy method for generating a population of templates for sequencing. The method is based on the random insertion of antibiotic resistance gene in plasmid DNA digested by DNase I. The advantages of this approach are the small quantity of DNA necessary for mutagenesis and the complete independence from the restriction map of the plasmid. DNase I digestion provides a random distribution of the insertions, while antibiotic selection provides low background. We also present a convenient PCR-based procedure for the analysis and ordering of obtained insertion mutants.

INTRODUCTION

Sequencing of DNA fragments that are not long enough to be used effectively in the shotgun method but are longer than one sequencing run is a common task in molecular biology. There are a number of methods to generate overlapping sequencing templates: (i) primer walking; (ii) shotgun method: random digestion and sequencing of short fragments (1,4); (iii) the generation of nested deletions by *ExoIII*/Mung Bean nuclease (6) or DNase I (2) digestion; and (iv) transposon-based in vitro systems of Tn5 (EF:TN; Epicentre, Madison, WI, USA), Tn7 (GPC; New England Biolabs, Beverly, MA, USA) or Mu (TGS; Finnzymes, Espoo, Finland).

The variety alone of approaches used might indicate that all these methods include some disadvantages. Primer walking, for example, is a costly, sequential, and, therefore, slow method. The shotgun approaches, although being very successful for genome sequencing projects, are inefficient for shorter DNA fragments because a very

uneven distribution of templates may occur. The *ExoIII*/Mung Bean method is a very labor-intensive approach that depends on a restriction map of the fragment. The plasmid DNA used for this particular method should be of high quality. Additionally, the length of deletions poorly correlates with the time of digestion.

MATERIALS AND METHODS

Kanamycin resistance gene (*Kan^R*) was cloned in pBluescript SK II(-)/*EcoRV* vector (Stratagene, La Jolla, CA, USA) after PCR amplification with flanking primers: 5'-ATGAGCTGATA-TCACAAATTC-3' and 5'-CTCGATA-TCTTCCTTTTCAATTC-3' (*EcoRV*

Figure 1. RIAR mutagenesis. To analyze the orientation and position of *Kan^R* insertion, each clone was tested in reactions with four different sets of primers. PCR fragments appearing during analysis are indicated in gray. The length of PCR products from the "positive" reactions (marked by "+") is further analyzed by electrophoresis. If the *Kan^R* gene is inserted outside the sequencing fragment, then both PCR products are longer than the insertion.

Short Technical Reports

sites introduced in primers are in bold). Zeocin resistance gene (*Zeo^R*) was excised from the plasmid pEM7/*Zeo* (Invitrogen, Carlsbad, CA, USA) with *Sac*I and *Eco*RI and inserted into pBluescript SK II(+)/(*Sac*I + *Eco*RI) vector. In the resulting construction, two *Eco*RV restriction sites surrounded the *Zeo^R* gene. Zeocin was selected as a rare antibiotic. Both antibiotic genes worked equivalently well. Because of the high price of zeocin, we preferably used kanamycin wherever possible; thus, in the scheme described below, we will refer to the gene of antibiotic resistance as kanamycin. For insertion mutagenesis, *Kan^R* and *Zeo^R* genes (1.15 and 0.47 kbp) were prepared at a concentration of 0.5 µg/µL by excision with *Eco*RV and gel electrophoresis purification.

The analytical digestion of plasmid DNA was performed for 1 h at 37°C in eight tubes with approximately 0.2 µg plasmid DNA in 20 µL buffer (10 mM Tris-HCl, pH 7.8, 1 mM MnCl₂, 0.1 µg/µL BSA) by serial two times dilutions of DNase I (RQ1 DNase; Promega, Madison, WI, USA). Mn²⁺-containing buffer was used for DNase I digestion to induce preferably double-strand breaks, but no nicks. The highest concentration of DNase I was 100 µU/reaction. The DNase I dilution buffer was (10 mM HEPES, pH 7.5, 10 mM CaCl₂, 10 mM MnCl₂, 100 µg/mL BSA, and 50% glycerol). Diluted DNase I was stable at -20°C for at least one year.

Preparative digestion was performed in one tube and differed from the analytical one only by blunt-ending of DNA. After 55 min at 37°C, the following components were added: 2.3 µL 10× T4 DNA polymerase buffer (500 mM NaCl, 100 mM Tris-HCl, pH 7.9, 100 mM MgCl₂, 10 mM DTT), 0.5 µL of mixture of four dNTPs (10 mM each), and 1 U T4 DNA polymerase (4). The mixture was incubated at 37°C for an additional 5 min.

Linear DNA was purified by gel electrophoresis and eluted in about 50 µL 15% PEG-TAE (5). Ligation was performed overnight at 14°C in 5 µL [3.5 µL eluted fragment (final concentration is about 3 ng/µL), 15 ng/µL *Kan^R* fragment, and 1 U T4 DNA ligase in appropriate buffer].

Bacteria with insertion of *Kan^R* gene were selected on ampicillin- (100 µg/

mL) and kanamycin- (50 µg/mL) containing media.

PCR for mapping the generated insertions was performed in 10–15 µL of a buffer containing 50 mM Tris-Cl, pH 8.6, 50 mM KCl, 1.5 mM MgCl₂, 0.1% Tween® 20, 50 µM Cresol red, 1.5 M betaine (3), 0.5 µg/mL ethidium bromide, 0.2 mM dNTPs, 0.1 µM two primers, and a mixture of *Taq* and *Pfu* DNA polymerases (100:1; 50 U/mL). Primers corresponding to the ends of the *Kan^R* gene were: (#1*k*) 5'-CGTTTCT-GCGGACTGGCTT-3' and (#2*k*) 5'-TC CCGATTTCGACGCGCAT-3'. Vector polylinker primers were: (#1*v*) M13 (-20) Forward 5'-TGTAACGACG GCCAGTGAATTGTAATAC-3' and (#2*v*) M13 Reverse 5'-TTGGCGTAAT-CATGGTCATAGCTGTTTCC-3'. The annealing temperature for reactions with these primers was 62°C. Betaine made it possible to load PCR mixture into the gel without loading buffer. Cresol red served as the gel running dye.

The amplification primers described above (#1*k* and #2*k*) or special sequencing primers (#1*k_s*) 5'-CTACGTGTTC-CGGTTCCTT-3' and (#2*k_s*) 5'-CATCGCCTTCTATCGCCTTC-3' may be used for sequencing of chosen clones. The primers were located at the following distances from the ends of the *Kan^R* fragment: 49 bp (#1*k*), 50 bp (#2*k*), 29 bp (#1*k_s*), and 33 bp (#2*k_s*).

RESULTS AND DISCUSSION

The insertion mutagenesis scheme of the random insertion of antibiotic re-

sistance (RIAR) method and the clone analysis are outlined in Figure 1.

The concentration of DNase I was adjusted to yield the highest amount of linear molecules in 1 h at 37°C (Figure 2). The low concentration of DNase I and the relatively long time of digestion were selected to convert the termini to blunt-ended during the last 5 min of digestion by fast-working T4 DNA polymerase in the same tube with no DNA purification.

For insertion mutagenesis of a number of different plasmids, we usually perform the analytical digestion only for one of them. In our practice, it is enough to take roughly the same quantity of other plasmid and estimated "optimal" concentration of DNase I to obtain the acceptable rate of digestion.

As shown in Figure 1, clones having the insertion of the antibiotic resistance gene outside the region to be sequenced can be identified during PCR analysis. The proportion of such clones may be roughly estimated as $l_v/(l_v + l_{ins})$, where l_{ins} is the length of the region to be sequenced and l_v is the length of the vector without antibiotic resistance gene and origin of replication. For a typical 3-kbp ($l_v \approx 1.26$ kbp) plasmid with 4-kbp insertions, the fraction of useless clones is about 24%. For 7 kbp, the fraction is about 15%. In principle, it is also possible that two antibiotic resistance genes will insert in one plasmid, but we never observed such clones.

The minimal number of colonies needed to sequence the insertion is l_{ins}/l_{read} , where l_{read} is the length of the sequencing run; 2.5–3 times more

Figure 2. Analytical digestion. Lanes 1–8 correspond to the sequential two times dilutions of DNase I. Following the increase of DNase I concentration, the band of supercoiled plasmid is transformed into separate, distinct bands of nicked circular and linear DNA. Original plasmid (lane 9) and the plasmid digested by a single cutting restriction enzyme (data not shown) can be used as controls of supercoiled and linear DNA. Lanes 2–4 correspond to the optimal digestion.

Figure 3. Examination of a 384-well PCR plate under UV illumination. Templates were loaded in the same order in four outlined zones. Primers used for each zone are indicated on the top panel. Crosses mark empty wells. As expected, only half of the PCR setups were successful.

than the minimal number of colonies was randomly taken for PCR analysis. Four reactions with different primer combinations were performed for each colony (Figure 1). To reveal the wells in which PCR took place, the PCR plate was examined under UV illumination (Figure 3).

The position of each insertion is determined by the length of two PCR fragments (Figure 4). An increase in the length of a particular PCR fragment is accompanied by a length reduction of the “second” PCR fragment. In case the insertion is located close to the ends of the sequenced fragment, it can be mapped more precisely according to the length of the shorter PCR fragment. When insertion is in the middle, it is possible to compare two PCR fragments with each other.

We have sequenced more than 25 kbp using the RIAR method, and the maximum length of individual template was

Short Technical Reports

Figure 4. Analysis of PCR fragments. PCR products from "positive" wells were separated in agarose gels. The white lines mark both PCR fragments from the same clone. The molecular weight marker is marked by asterisks. (A) Preliminary analysis. (B) Second control electrophoresis of the selected clones ranged according to the proposed insertion positions to compare the neighboring insertions more precisely. Photos for three different plasmids are presented. The decrease of the length of one PCR fragment combined with a corresponding increase of the matching "second" fragment results in a typical butterfly-like picture on the gel.

about 6 kbp. The mean length of gaps after DNase I digestion and blunt-ending was about 20 bp (data not shown).

Compared to existing nuclease-based approaches (2,6), the RIAR method is less time and material consuming, independent from the restriction map of the plasmid and robust to variations in plasmid DNA quality (nicks and RNA contamination did not interfere with the method). The distribution of insertions depends on DNase I digestion and is random enough for most applications. Each insertion allows sequencing in both directions; thus, the necessary number of plasmids is two times lower than in nested deletion methods.

Transposon-based in vitro cloning systems are slightly easier to use. Their other advantage is that transposon insertion induces duplication of several nucleotides adjacent to the insertion site. It enables an unambiguous assembly of two sequences generated from each end of insertion. Compared to transposon-based systems, the RIAR

system has a lower price [for transposon-based systems, GPC-1 (New England Biolabs), it is \$250/10 reactions], and it is possible to adapt it easily to a variety of plasmids and antibiotics. The PCR-based screening method presented here may also be useful for the transposon-based in vitro cloning systems.

REFERENCES

1. Demolis, N., L. Mallet, F. Bussereau, and M. Jacquet. 1995. Improved strategy for large-scale DNA sequencing using DNase I cleavage for generating random subclones. *BioTechniques* 18:453-457.
2. Eberle, J.R. 1993. Generation of bidirectional deletions using DNase I. *BioTechniques* 14:408-411.
3. Henke, W., K. Herdel, K. Jung, D. Schnorr, and S.A. Loening. 1997. Betaine improves the PCR amplification of GC-rich DNA sequences. *Nucleic Acids Res.* 25:3957-3958.
4. Sambrook, J. and D. Russell. 2001. *Molecular Cloning: A Laboratory Manual*. CSH Laboratory Press, Cold Spring Harbor, NY.
5. Zhen, L. and R.T. Swank. 1993. A simple and high-yield method for recovering DNA from agarose gels. *BioTechniques* 14:894-898.
6. Zhu, K.Y. and J.M. Clark. 1995. Rapid construction of nested deletions of recombinant plasmid DNA for dideoxy sequencing. *BioTechniques* 18:222-224.

The authors are greatly indebted to H. Eickhoff for help with the preparation of the text of the paper. The work of E.N. and S.G. was supported by a fellowship from Oslo University; EMBO fellowship ALTF517-1999 and BMBF grant 0312275F/2 to A.S.; grant N 00-04-22000 from the Russian Foundation for Basic Research to E.N., S.G., and A.K. Address correspondence to Dr. Aleksey Soldatov, Max-Planck-Institut für Molekulare Genetik, Abteilung Lehrach, Ihnestrasse 73, 14195 Berlin-Dahlem, Germany. e-mail: soldatov@molgen.mpg.de

Received 9 July 2001; accepted 13 September 2001.

**Elena N. Nabirochkina,
Sofia G. Georgieva, Aleksey
N. Krasnov, and Aleksey
V. Soldatov**
*Institute of Gene Biology
RAS Moscow
University of Oslo
Center for Medical Studies
Moscow, Russia
Max-Planck-Institut für
Molekulare Genetik
Berlin, Germany*

**For reprints of this or
any other article, contact
Reprints@BioTechniques.com**