Generation of Genomic Mini-Libraries by *Taq* DNA Polymerase Modification of Genomic Fragments

BioTechniques 24:566-568 (April 1998)

When cloning genomic DNA fragments, the efficiency is greatly reduced as the size of the restriction fragment increases. Libraries derived from the total genome are therefore biased because shorter fragments will be cloned preferentially. When cloning a particular gene, it is most desirable to determine a restriction enzyme that will generate a fragment of a reasonable size for the gene of interest by Southern blotting. A mini-library can then be created containing only the genomic fragment of that particular size and reducing the number of colonies that have to be screened. This procedure will inevitably mean that a different restriction enzyme is used for each fragment cloned, and therefore a fresh batch of digested, phosphatased vector must be prepared each time. The digestion and dephosphorylation reactions must both be highly efficient to obtain a high percentage of insert uptake, and often this is the rate-limiting step in mini-library creation. We have circumvented this problem by using a commercially available vector designed for the cloning of polymerase chain reaction (PCR) products. The pGEM[®]-T Vector (Promega, Madison, WI, USA) contains a single 3' T overhang that allows the insertion of PCR products with overhanging A residues added by the terminal transferase activity of certain *Taq* DNA polymerases that lack proofreading activity (1,2). *Taq* DNA polymerase can be used to create similar overhanging ends on any blunt-ended fragment, and these can then be efficiently cloned into pGEM-T, which gives a low background of recircularization without insert. This method also increases the range of restriction enzymes that can be screened to identify a fragment of the desired size because it is not limited by those present in vector multiple cloning sites. It is equally useful if two restriction enzymes are used to generate a

fragment of the desired size.

This method is particularly useful in isolating clones from organisms whose DNA is of an unusual nucleotide content and so tends to be unstable when propagated in bacteria. For example, the genomes of *Dictyostelium discoideum* and *Malaria plasmodium* are both extremely AT-rich and notoriously unstable, especially in high-copy-number vectors. This means that it is impractical to amplify a genomic library for repeated use because the amplification procedure leads to loss and rearrangement of a significant percentage of the clones. Thus, even if many genes of interest can in theory be isolated from the same library, in practice it is necessary to regularly remake the library to avoid amplification steps.

We have used this system to clone a 3-kb genomic fragment containing the *mo15* gene from the cellular slime mold *D. discoideum*. Genomic clones were identified by hybridization to a partial cDNA clone from *Dictyostelium* encoding a protein highly homologous to *Xenopus* cyclin-dependent protein kinase MO15 (cdk7) (3). As stated above, isolating genomic clones from this particular organism can be problematic because the extremely AT-rich nature of the DNA makes many fragments unstable for cloning into high copy-number vectors. We have made several unsuccessful attempts to isolate genomic clones encoding *mo15* from conventional *Dictyostelium* libraries that had successfully been used to isolate other genes.

Purification of 3-kb fragments from *EcoRI*-digested *Dictyostelium* genomic DNA. Southern blot analysis of *Dictyostelium* genomic DNA digested with a range of restriction enzymes revealed that the partial *mo15* cDNA clone available hybridized to a 3-kb *EcoRI* fragment (data not shown). Because the coding region of the *mo15* gene is believed to be around 1.2 kb and there is no *EcoRI* site within this, it was decided to isolate this 3-kb fragment because it was likely to contain the entire *mo15* gene. Aliquots of restriction digest (20 and 60 µg) were loaded onto a 0.8% agarose gel, and after resolving the fragments by electrophoresis, the region around 3 kb was excised from the gel. The DNA was eluted using the GENECLEAN[®] II Kit

(Bio 101, Vista, CA, USA) according to the manufacturer's instructions.

Cloning of fragments into pGEM-T. The purified fragments were incubated with 25 mM dNTPs and 5 U of *Taq* DNA polymerase (Boehringer Mannheim, Lewes, East Sussex, England, UK) in the buffer recommended by the manufacturer at 72°C for 30 min. The polymerase activity of the *Taq* enzyme created blunt-ended fragments, and then the terminal transferase activity caused the addition of A residues to the 3' ends of the blunt fragments. After cleaning with GENECLEAN II, the fragments were ligated with 25 ng of pGEM-T at 4°C for 18 h in a 10-µL ligation mixture. This mixture (2.5 µL) was transfected into UltraMAX DH5α[™] Competent Cells (Life Technologies, Paisley, Scotland, UK) by a chemical transformation method and 10% of the cells plated onto ampicillin (100 µg/mL) resistance plates containing 80 µg/mL 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-gal) and 0.5 mM isopropyl β-D-thiogalactopyranoside (IPTG) to allow for blue/white selection of the colonies with inserts. The percentage of white colonies containing insert was dependent on the amount of genomic DNA originally present in the restriction digest. In this instance, starting with 60 µg of genomic DNA rather than 20 µg led to a 9-fold increase in the number of white colo-

Figure 1. Average insert size in genomic mini-library. Six white colonies were picked at random and the plasmid DNA isolated using the Wizard[®] purification procedure (Promega). The DNA was digested with *PvuII*, which removes the insert and 500 bp of flanking vector DNA, and resolved by electrophoresis on a 0.8% agarose gel. The Ladder Mix (MBI Fermentas Ltd., Vilnius, Lithuania) was run to allow size determination (M). The vector (2567 bp) and a single insert band are seen in all six cases.

Benchmarks

nies and a corresponding reduction in the number of blue colonies by 50%. The optimum amount of genomic DNA will vary for each mini-library created because it is dependent on the number of restriction fragments of that size present in the genome of interest and must be determined empirically.

The average insert size was checked in six colonies picked at random (Figure 1). The DNA was digested with *PvuII*, which removes the insert and 500 bp of flanking vector DNA. Of the six colonies, four have inserts of around 3.0–3.5 kb, as expected. Two clones contained much smaller fragments of DNA. This could reflect contamination of the size-selected insert in combination with preferential insertion of smaller fragments. Alternatively, it could represent instability of the AT-rich *Dictyostelium* DNA when present in bacteria. We favor the second possibility because this is a known feature of *Dictyostelium* genomic DNA.

Isolation of the *mo15* gene. Colonies derived from the the most efficient insertion ligation were screened using the *mo15* cDNA clone as probe (3), and out of approximately 8000 white colonies screened, 20 came through two rounds of screening containing the *mo15* gene. This was confirmed by PCR analysis using primers based on the known coding sequence (data not shown). When DNA was isolated from two of these, the insert size was found to be approximately 3 kb, as expected, and DNA sequence analysis confirmed the presence of the *mo15* gene. These are independent isolates because restriction analysis revealed that each was in a different orientation relative to the vector sequences (data not shown).

This quick and efficient method for mini-library creation increases the range of restriction enzymes that can be used to isolate a fragment of suitable size and circumvents the necessity to generate a batch of digested, phosphatased vector with a low background of self-ligation for each restriction enzyme used. It is particularly useful when isolating clones from organisms whose DNA is unstable in bacteria, because the desire not to amplify libraries means that it is often necessary to generate a fresh library several times to isolate more than one gene.

REFERENCES

1. Clark, J.M. 1988. Novel non-templated nucleotide addition reactions catalyzed by prokaryotic and eucaryotic DNA polymerases. *Nucleic Acids Res.* 16:9677.
2. Marchuk, D., M. Drumm, A. Saulino and F. Collins. 1991. Construction of T-vectors, a rapid and general system for direct cloning of unmodified PCR products. *Nucleic Acids Res.* 19:1154.
3. Michaelis, C., Q. Luo and G. Weeks. 1995. A *Dictyostelium discoideum* gene which is highly related to *mo15* from *Xenopus* is expressed during growth but not development. *Biochem. Cell Biol.* 73:51–58.

*The authors would like to acknowledge the support of a Royal Society Equipment grant and Wellcome Grant No. 039430. H.-J.H. is the grateful recipient of an ORS award. Many thanks to Wen-Tsang Chang for helpful comments on reading the manuscript and to Gerry Weeks for providing the partial *mo15* cDNA clone. Address correspondence to Catherine Pears, Department of Biochemistry, University of Oxford, South Parks Road, Oxford, OX1 3QU, England, UK. Internet: pears@bioch.ox.ac.uk*

Received 22 August 1997; accepted 18 November 1997.

Hao-Jen Huang and Catherine Pears

*University of Oxford
Oxford, England, UK*

Small-Scale Preparation of the Single-Copy Bacterial Artificial Chromosome Vector pBeloBAC11

BioTechniques 24:568–572 (April 1998)

The construction of stable, large-insert DNA libraries is critical for the analysis of complex genomes (8). Traditionally, yeast artificial chromosomes (YACs) (1) have been the system of choice, but this system suffers from a number of disadvantages. YAC libraries can be difficult and tedious to construct, cloned sequences are often unstable or