Short Technical Reports

5. Hanazono, Y., J.M. Yu, C.E. Dunbar and R.V. Emmons. 1997. Green fluorescent protein retroviral vectors: low titer and high recombination frequency suggest a selective disadvantage. *Gene Ther.* 8:1313-1319.
6. Lybarger, L., D. Dempsey, K.J. Franek and R. Chervenak. 1996. Rapid generation and flow cytometric analysis of stable GFP-expressing cells. *Cytometry* 25:211-220.
7. Miller, J.L., C.E. Walsh, P.A. Ney, R.J. Samulski and A.W. Nienhuis. 1993. Single-copy transduction and expression of human gamma-globin in K562 erythroleukemia cells using recombinant adeno-associated virus vectors: the effect of mutations in NF-E2 and GATA-1 binding motifs within the hypersensitivity site 2 enhancer. *Blood* 82:1900-1906.
8. Misteli, T. and D.L. Spector. 1997. Applications of the green fluorescent protein in cell biology and biotechnology. *Nat. Biotechnol.* 15:961-964.
9. Niswender, K.D., S.M. Blackman, L. Rohde, M.A. Magnuson and D.W. Piston. 1995. Quantitative imaging of green fluorescent protein in cultured cells: comparison of microscopic techniques, use in fusion proteins and detection limits. *J. Microsc.* 180:109-116.
10. Okabe, M., M. Ikawa, K. Kominami, T. Nakanishi and Y. Nishimune. 1997. 'Green mice' as a source of ubiquitous green cells. *FEBS Lett.* 407:313-319.
11. Persons, D.A., J.A. Allay, E.R. Allay, R.J. Smeyne, R.A. Ashmun, B.P. Sorrentino and A.W. Nienhuis. 1997. Retroviral-mediated transfer of the green fluorescent protein gene into murine hematopoietic cells facilitates scoring and selection of transduced progenitors in vitro and identification of genetically modified cells in vivo. *Blood* 90:1777-1786.
12. Riera, F.C., S.B. Blam, N.N. Teng and H.S. Kaplan. 1984. Somatic cell hybrid selection with a transfectable dominant marker. *Somat. Cell Mol. Genet.* 10:123-127.
13. Sasaki, K., H. Mizusawa, M. Ishidate and N. Tanaka. 1992. Regulation of G418 selection efficiency by cell-cell interaction in transfection. *Somat. Cell Mol. Genet.* 18:517-527.
14. Schott, B., E.S. Iraj and I.B. Roninson. 1996. Effects of infection rate and selection pressure on gene expression from an internal promoter of a double gene retroviral vector. *Somat. Cell Mol. Genet.* 22:291-309.
15. Somlyai, G., E. Kondorosi, K. Kariko and E.G. Duda. 1985. Liposome mediated DNA-transfer into mammalian cells. *Acta Biochim. Biophys. Acad. Sci. Hung.* 20:203-211.
16. Southern, P.J. and P. Berg. 1982. Transformation of mammalian cells to antibiotic resistance with a bacterial gene under control of the SV40 early region promoter. *Mol. Appl. Genet.* 1:327-341.
17. Sutherland, H.G., D.I. Martin and E. Whitelaw. 1997. A globin enhancer acts by increasing the proportion of erythrocytes expressing a linked transgene. *Mol. Cell Biol.* 17:1607-1614.
18. Welsh, S. and S.A. Kay. 1997. Reporter gene expression for monitoring gene transfer. *Curr. Opin. Biotechnol.* 8:617-622.
19. Zlokarnik, G., P.A. Negulescu, T.E. Knapp, L. Mere, N. Burres, L. Feng, M. Whitney, K. Roemer and R.Y. Tsien. 1998. Quantita-

tion of transcription and clonal selection of single living cells with beta-lactamase as reporter. *Science* 279:84-88.

Address correspondence to Jeffery L. Miller, Laboratory of Chemical Biology, Building 10, Room 9N308, National Institutes of Health, Bethesda, MD 20892, USA. Internet: jm7f@nih.gov

Received 6 May 1999; accepted 3 August 1999.

A.N. Gubin, S. Koduru, J.M. Njoroge, R. Bhatnagar and J.L. Miller
*National Institute of Diabetes and Digestive and Kidney Diseases
National Institutes of Health
Bethesda, MD, USA*

Restriction Cutting Independent Method for Cloning Genomic DNA Segments Outside the Boundaries of Known Sequences

BioTechniques 27:1170-1177 (December 1999)

ABSTRACT

We present a simple method for cloning genomic DNA segments outside the boundaries of known sequences, which is not dependent on restriction cutting or mapping. In the first step of the method, a library of single-stranded flanking sequences is generated by linear amplification with one primer in the known region. A homooligomeric cytosine tail is added to each of the single-stranded fragments by a terminal transferase catalyzed reaction. The tailed fragments are then amplified by PCR with a nested primer in the known region and a poly-guanine primer complementary to the cytosine tail in the unknown region. Finally, the different fragments are separated by cloning and characterized by sequencing. The method was used to clone both the upstream (5') and the downstream (3') genomic regions of an intron-interrupted *tRNA^{Leu}(UAA)* gene from three cyanobacteria belonging to the genus *Microcystis*.

INTRODUCTION

Traditionally, cloning of regions outside the boundaries of known sequences is dependent on the screening of genomic libraries with ³²P-labeled probes (13). Several PCR-based approaches, such as inverse PCR (9) and adapter-mediated techniques (2,3,5,7, 11), have emerged over the last decade. Although these techniques in many respects have simplified the cloning of flanking sequences, they still depend on a relatively labor-intensive restriction cutting for finding suitable restriction enzymes. These methods are also limited by their robustness and do not always result in successful cloning of flanking sequences.

We present a new and simple

method for genomic flank cloning from genomes with low complexities. The method is not dependent on restriction enzyme cutting or mapping. The method uses the same homo-oligomeric tailing principle as for rapid amplification of cDNA ends (RACE) (1,4,10). However, the fundamental difference from RACE is that the described method is designed for cloning genomic flanking regions and not cDNA ends. The described method is based on generating single-stranded (ss)DNA covering the flanking region by linear amplification with one primer in the known region. The different fragments are tailed with cytosine by a terminal transferase catalyzed reaction. Then, the tailed fragments are amplified by PCR with a nested primer in the known region and a homo-oligomeric poly-guanine primer complementary to the cytosine tail in the unknown region. Finally, the different fragments are cloned by TA cloning (6) and characterized by sequencing. In this work, the versatility of the method was demonstrated by successful cloning of both the upstream and downstream regions of an intron-interrupted tRNA^{L^{eu}}(UAA) gene in 3 cyanobacteria belonging to the genus *Microcystis* (12).

MATERIALS AND METHODS

Sample Preparation

The strains *Microcystis aeruginosa* NIVA-CYA 143, 57 and 228/1 were cultivated in glass flasks containing the medium Z8 in a constant temperature room at $17 \pm 2^\circ\text{C}$ and exposed to light in the range 10–60 microeinsteins $\text{m}^{-2} \text{s}^{-1}$ (8). Aliquots of dense culture (50 mL containing approximately 10^7 cells/mL) were pelleted in a GLC-1 swing-out centrifuge (Kendro Laboratory Products, Newtown, CT, USA) at 3000 rpm for 15 min. The DNA was then purified from the concentrated cells by a standard phenol/chloroform method (13).

Cloning and Sequencing of Flanking Regions

Primers 5'-TGGGGGTGGGGG-
ACTTGAAC-3' (CB; for cloning of the

upstream region) and 5'-CGCCAGCG-
GGGCACGGTGA-3' (EC; for cloning
the downstream region) were used in
the linear amplification reaction (Fig-
ure 1A and for location of the primers,
see Figure 2). The 200 μL reaction con-
tained 2 pmol primer, 200 μM dNTP,
10 mM Tris-HCl (pH 8.8), 1.5 mM
 MgCl_2 , 50 mM of KCl, 0.1% Triton®
X-100, 3 U of DynaZyme™ Thermo-
stable DNA Polymerase (Finnzymes
Oy, Espoo, Finland) and approximately
0.5 μg DNA. The linear amplification
was performed with 25 cycles, using
the following cycling parameters: 30 s
at 94°C , 30 s at 50°C or 60°C (for
primer CB and EC, respectively) and fi-
nally, 1 min at 72°C . The DNA was de-
natured for 4 min at 94°C before the
amplification.

For the cytosine tailing reaction
(Figure 1B), all of the linear amp-
lification products were extracted once
with phenol/chloroform/isoamyl alco-
hol (25:24:1) and once with chloro-
form/isoamyl alcohol (24:1). The DNA
was then precipitated with 2 vol of
ethanol, 0.1 vol of 3M sodium acetate
(pH 5.2) and 0.01 vol of 2 mg/mL
glycogen (Roche Molecular Biochemi-
cal, Mannheim, Germany) at -20°C for
2 h. The precipitated DNA was pelleted
by centrifugation at $15\,000\times g$ in a mi-
crocentrifuge at 4°C for 30 min. The
pellet was rinsed with 70% ethanol,
dried in a vacuum centrifuge and resus-
pended in 10 μL of water. The tailing
reaction contained 5 μL of the purified
amplification products, 25 U of calf
thymus nucleoside-triphosphate: DNA
deoxynucleotidylexotransferase (Roche
Molecular Biochemical), $1\times$ reaction
buffer for terminal transferase contain-
ing 0.2 M potassium cacodylate, 25
mM Tris-HCl, 0.25 mg/mL bovine
serum albumin and 0.75 mM CoCl_2 in
a final vol of 10 μL . The reaction mix-
ture was incubated at 37°C for 20 min
and then at 96°C for 4 min to inactivate
the enzyme.

The nested primers 5'-GAACCC-
CAAGGCAAAGCCCG-3' (EB) and
5'-TCGGGCWACCGAGTGGCGG-3'
(ED; for cloning of the upstream and
downstream regions, respectively; for
location of the primers see Figure 2),
and the poly-guanine primer 5'-AT-
ATGGGGGGGGGGGGGGG-3' (EA)
were used in the PCR amplification for

Short Technical Reports

generating a library of fragments covering the flanking regions (Figure 1C). The 50 μ L reaction contained 10 pmol of each of the primers and 2 μ L of the tailed fragments. The other reaction components were the same as for the linear amplification. The cycling was done for 35 cycles, using the following cycling parameters: 96°C for 30 s, 65°C for 30 s and 72°C for 2 min. The reaction was started with a denaturation

at 94°C for 4 min and ended with an extension for 7 min at 72°C.

The amplification products were purified by using the QIAquick™ PCR Purification Kit (QIAGEN, Hilden, Germany) as recommended by the manufacturer. The amplified products were cloned, using the pGEM®-T Vector System (Promega, Madison, WI, USA) according to the manufacturer's recommendations. Finally, the flanking

sequences were sequenced automatically on a Model 377 DNA Sequencer (PE Biosystems, Foster City, CA, USA) using the dye terminator sequencing chemistry, as recommended by the manufacturer.

RESULTS AND DISCUSSION

The objective of the work presented here was to develop a simple method for cloning genomic flanking regions, avoiding restriction cutting and mapping. This was achieved by homopolymer tailing of a population of single-stranded amplification products covering the flanking region. The designed approach was tested for rapid screening of tRNA flanking regions from cyanobacteria.

Generation of Libraries of tRNA^{leu}(UAA) Flanking Sequences

The cloning of the flanking regions of an intron-interrupted tRNA^{leu}(UAA) gene in the cyanobacteria *M. aeruginosa* strains NIVA-CYA 143, 57 and 228/1 were used as model systems. The flanks were cloned by use of the described method for NIVA-CYA 57, while the other flanks were identified through homology to NIVA-CYA 57. Both the upstream (5') and the downstream (3') regions were cloned in all cases.

In a one-sided amplification, fragments of different sizes were generated by using tRNA^{leu}(UAA)-specific primers CB and EC for the upstream and downstream regions, respectively (Figure 2). For the control where this step was omitted, no flanking sequences were retrieved. By empirical screening, the optimal number of cycles was deter-

Figure 1. Schematic representation of the method. (A) A library of single-stranded fragments covering the unknown region was generated by a linear amplification with a primer in the known region. (B) A homo-oligomeric cytosine tail was then added by a terminal transferase catalyzed reaction. (C) Finally, the tailed fragments were amplified with a poly-guanine primer complementary to the cytosine tail in the unknown region and a nested primer in the known region.

Figure 2. Flanking region and primer sites for *M. aeruginosa* NIVA-CYA 57. Two open reading frames (ORF 1 and 2) were identified in the flanking regions. The primer sites and orientations are indicated with arrows for primer CB, EB, EC and ED. The scaling are base pairs starting from the region upstream to the tRNA gene.

Short Technical Reports

mined to be 25. This number is a trade off between sensitivity and the generation of nonspecific amplification products. The fragments from the linear amplification were subsequently used for tailing with cytosine and amplified by EB and ED (upstream and downstream, respectively) (Figure 2) and EA. A smear of fragments with different sizes was generated after the second amplification (Figure 3). Finally, the different fragments were cloned by TA cloning and characterized by sequencing.

Sequence Analysis of the Flanking Regions

As determined by sequencing, approximately 44% of the clones had inserts containing flanking regions (4 out of 9) with an average size of 1068 bp for the upstream region, while for the downstream region, 20% of the clones

had flank inserts (1 out of 5) with a size of 250 bp. The longest fragment retrieved in our experiments was 1312 bp. The lower efficiency for the downstream region, as compared to the upstream region, may be due to differences in the properties of the primers used. However, the potential of the method is demonstrated through the cloning of the upstream region. A PCR primer pair complementary to the 5' and 3' flanking regions was constructed to test whether the identified regions flank the same tRNA^{Leu}(UAA) gene. A single amplification product was obtained with this primer pair for all the strains tested (results not shown). It was also confirmed by DNA sequencing that the cloned regions flank the same tRNA^{Leu}(UAA) gene. Interestingly, the nonspecific amplification products were in all cases results of amplification with EB or ED and not EA.

Comparison with Other Methods for Cloning Flanking Regions

Flank cloning strategies are generally dependent on restriction cutting or mapping for identifying suitable restriction sites and enzymes. This process alone can take up to several weeks for each flank. The strategy presented here is not dependent on restriction enzyme cutting, thus simplifying the process. Flanks can be cloned solitarily based on the sequence information in the known region. The whole process takes about 24 h, with approximately 4 h hands-on time. The simplicity of the process is important for comparative flank studies, for example. Preliminary comparative analyses suggest that the tRNA^{Leu}(UAA) gene is located in a conserved region of the genome (Figure 2 and unpublished data).

Because relatively crude DNA

Figure 3. Nested PCR using internal (nested) primers. The nested amplification products were obtained by amplification with the primer pair EA and EB, as described in the text. Ten microliters of the products were electrophoresed in an ethidium bromide-stained 1.5% agarose gel for 45 min at 100 V. The lane labeled "pos" contains the reaction with template (seen as a weak smear), while the lane labeled "neg" contains the reaction without template. mw, molecular weight marker.

preparations can be used in amplification reactions, using thermal cycling, the strategy may also be applied on samples that are too crude for other methods. Finally, the method is not dependent on double-stranded genomic DNA, so that DNA denatured by heat inactivation of enzymes, for example, can be used for cloning flanking sequences. Taken together, the method is likely to represent a major improvement for the characterization of genomic flanking regions.

Other Applications of the Cloning Strategy

Flanking fragments of more than 1000 bp were easily retrieved by the protocol applied here. However, larger fragments can probably be cloned through size separation (e.g., by gel electrophoresis before the construction of the random library). Cloning of larger fragments is preferable for many applications, such as primer walking.

Cloning of flanking sequences from complex genomes may also be an ap-

plication of the technique. Using two rounds of nested amplification will probably reduce the background of nonspecific amplification products considerably. (Our experience is that the polyguanine primer EA gave very low background of nonspecific products). Work is in progress adapting the technique for cloning of flanking sequences from organisms with complex genomes, such as higher eukaryotes.

REFERENCES

1. Chenchik, A., L. Diachenko, F. Moqadam, V. Tarabykin, S. Lukyanov and P.D. Siebert. 1996. Full-length cDNA cloning and determination of mRNA 5' and 3' ends by amplification of adapter-ligated cDNA. *BioTechniques* 21:526-534.
2. Cormack, R.S. and I.E. Somssich. 1997. Rapid amplification of genomic ends (RAGE) as a simple method to clone flanking genomic DNA. *Gene* 194:273-276.
3. Espelund, M. and K.S. Jakobsen. 1992. Cloning and direct sequencing of plant promoters using primer-adapter mediated PCR on DNA coupled to a magnetic solid phase. *BioTechniques* 13:74-81.
4. Gurr, S. and M. McPherson. 1991. PCR-directed cDNA libraries. In M. McPherson, P. Quirke and G. Taylor (Eds.), *PCR a Practical Approach*, Oxford University Press.
5. Iwahana, H. and M. Itakura. 1997. An end trimming method and its application to amplify adjacent cDNA and genomic DNA fragments by PCR. *Methods Mol. Biol.* 67:247-260.
6. Marchuk, D., M. Drumm, A. Saulino and F.S. Collins. 1991. Construction of T-vectors, a rapid and general system for direct cloning of unmodified PCR products. *Nucleic Acids Res.* 19:1154.
7. Moynihan, T.P., A.F. Markham and P.A. Robinson. 1996. Genomic analysis of human multigene families using chromosome-specific vectorette PCR. *Nucleic Acids Res.* 24:4094-4095.
8. Norwegian Institute for Water Research. 1990. Culture Collection of Algae. Catalogue of Strains. Norwegian Institute for Water Research, Oslo, Norway.
9. Ochman, H., A.S. Gerber and D.L. Hartl. 1988. Genetic applications of an inverse polymerase chain reaction. *Genetics* 120:621-623.
10. Ohara, O., R.L. Dorit and W. Gilbert. 1989. One-sided polymerase chain reaction: the amplification of cDNA. *Proc. Natl. Acad. Sci. USA* 86:5673-5677.
11. Rosenthal, A. and D.S. Jones. 1990. Genomic walking and sequencing by oligo-cassette mediated polymerase chain reaction. *Nucleic Acids Res.* 18:3095-3096.
12. Rudi, K. and K.S. Jakobsen. 1997. Cyanobacterial tRNA^(Leu)(UAA) group I introns have polyphyletic origin. *FEMS Microbiol. Lett.* 156:293-298.
13. Sambrook, J., E.F. Fritsch and T. Maniatis.

1989. *Molecular Cloning: A Laboratory Manual*, 2nd ed. CSH Laboratory Press, Cold Spring Harbor, New York.

This work was supported by Grant No. 118894/431 from the Norwegian Research Council (NFR) to K.S.J. We thank Olav M. and Randi Skulberg for the Microcystis cultures, Odd-Arne Olsen for the use of the PE Biosystems Model 377 DNA sequencer, Hege Munch for running the sequence reactions and Heidi Rudi and John E. Stacy for critically reading and commenting on the manuscript. Address correspondence to Knut Rudi, MATFORSK Norwegian Food Research Institute, Osloveien 1, 1430 Ås, Norway. Internet: knut.rudi@matforsk.no

Received 22 January 1999; accepted 29 July 1999.

Knut Rudi, Tonje Fossheim and Kjetill S. Jakobsen
University of Oslo
Oslo, Norway