Efficient Cloning of SAGE Tags by Blunt-End Ligation of Polished Concatemers

BioTechniques 34:__-__ (April 2003)

The serial analysis of gene expression (SAGE) has become an important technique for high-throughput transcriptome analysis. Since its first description in 1995 (1), SAGE has provided both quantitative and qualitative information concerning global gene expression in a variety of biological systems. Thereby, it has deepened our knowledge concerning developmental processes as well as pathogenic mechanisms (2–4). Furthermore, SAGE was shown to be a useful tool for genome annotation (5).

The common application of SAGE, however, was delayed and is still hampered by a technically demanding protocol. Recently, improvements have been made toward a more widespread use of this method (6–9). Several of the methodical advances concerned the cloning of the concatemers, as this was shown to be a crucial yet inefficient step. One protocol described the removal of biotinylated linkers by streptavidin-coated magnetic beads before cloning (10). The introduction of a heating step before separation of the concatemers by gel electrophoresis resulted in a better correlation between the migration distance in the gel and the concatemer length (11). This allowed the cloning of concatemers with, on average, significantly greater length. None of these protocols, however, has addressed the major problem underlying inefficient cloning of any given concatemer, which is the problem of "corrupted concatemer termini". Corrupted concatemer termini do not contain the nucleotide termini expected from normal endonuclease cleavage; consequently, concatenation has an inherent tendency to stop at these termini, as they prevent further elongation. Corrupted termini can arise from exonuclease activity present during/after restriction endonuclease cleavage, or from star activity of the restriction endonuclease used. These corrupted termini result in poor cloning efficiency of the concatemers, as these are subsequently cloned into a vector cleaved with a restriction endonuclease that yields cohesive ends compatible only with intact termini.

Table 1. Cloning Efficiency

Ligation Reaction	Transformants/ 1.5 µL Ligation	Total Number of Transformants[a]	%[b]
Fraction I, *Sph*I	7260	96 800	100
Fraction I, *Eco*32I	25 245	336 600	347
Fraction II, *Sph*I	3960	52 800	100
Fraction II, *Eco*32I	8085	107 800	204
pBluescript, *Sph*I[c]	9	n.d.	n.d.
pBluescript, *Eco*32I[c]	330	n.d.	n.d.

[a]Calculated number of transformants using the complete gel-eluted concatemers of a given size fraction.
[b]The clone number of the corresponding *Sph*I SAGE library was set to 100%.
[c]Self ligation of restriction endonuclease-cleaved and dephosphorylated vector.
n.d., not determined.

One approach to overcome this problem involves polishing the concatemer termini, thus allowing blunt-end ligation. Although blunt-end ligation is generally less efficient than cohesive end ligation, it can result in a higher efficiency under conditions of enriched corrupted cohesive termini. To test this hypothesis, we used the exonuclease activity of bacteriophage T4 DNA polymerase to convert the 3′-protruding concatemer termini to blunt-end DNA, as the commonly used fill-in reaction by Klenow is prohibited by the presence of protruding 3′ ends (12).

Concatemers were purified from biotinylated linkers by streptavidin-coated magnetic beads (Dynal) (10), kept at 65°C for 5 min to break concatemer aggregates (11) and separated in a 1% agarose gel (universal agarose; Peqlab). Fractions of 300–500 bp (fraction I) and 500–1000 bp (fraction II) in length were cut out and eluted from the gel matrix with the E.Z.N.A. Gel Extraction Kit (Peqlab). To compare blunt-end ligation with the established protocol of cloning concatemers into the *Sph*I restriction site, each fraction was divided into two aliquots of 15 µL. To one aliquot, 1.2 U T4 DNA polymerase (New England Biolabs, Beverly, MA, USA) were added in the presence of 100 µM each dNTP, and the reaction was run for 15 min at 12°C in 1× reaction buffer in a total volume of 20 µL. The reaction was stopped by heat-denaturing the enzyme at 75°C for 20 min. This yielded a polished concatemer sample ready for blunt-end ligation. The other aliquot was kept at 4°C during this procedure (unpolished concatemer sample).

To clone concatemers, a modified pBluescript® vector (Stratagene, La Jolla, CA, USA) was used that contained an additional *Sph*I recognition site between the *Bam*HI and the *Eco*RI recognition site. Five micrograms of vector were cleaved either with 50 U *Sph*I (New England Biolabs) or with 50 U *Eco*32I (MBI Fermentas) for 90 min at 37°C in a total volume of 100 µL. *Eco*32I represents an isoschizomer of *Eco*RV and was recently reported to improve blunt-end cloning by a decreased intrinsic single nucleotide deletion activity when compared to *Eco*RV (13). This should allow for more efficient blue/white selection. After cleavage by restriction endonuclease, vectors were dephosphorylated using 10 U calf intestinal phosphatase (New England Biolabs), purified by agarose gel electrophoresis, and extracted using the same gel extraction kit as above (Peqlab). The entire unpolished samples were ligated into 100 ng *Sph*I-cleaved pBluescript; the entire T4 DNA polymerase-polished samples were directly ligated after the heat denaturation step into 100 ng *Eco*32I-cleaved pBluescript by adding 15 U T4 DNA ligase (MBI Fermentas), 5 µL 10 × ligase buffer and water to a total volume of 50 µL. Ligations were carried out at 15°C for 19 h. Thereafter, T4 DNA ligase

was removed by StrataClean™ resin (Stratagene), and the ligation products were precipitated with ethanol in the presence of 70 µg glycogen (Peqlab) and resuspended in 10 µL water. Then, 1.5 µL each ligation product was added to 50 µL ElectroTen-Blue™ electroporation competent cells (≥ 1×10^{10} transformants/µg; Stratagene). After electroporation at 1.8 kV (*E. coli* pulser; Bio-Rad Laboratories, Hercules, CA, USA) in 0.2-cm-wide cuvettes (Equibio), transformation mixtures were spread on 145-mm-wide LB agar plates containing 50 µg/mL ampicillin (Sigma) and the X-gal/IPTG system (Peqlab) for blue/white selection. Transformed cells were counted the next day. A substantial increase of transformants was observed for T4 DNA polymerase-treated concatemers when compared to the corresponding *Sph*I ligation (Table 1). A 3.4-fold increase for fraction I and a 2-fold increase for fraction II were observed. On all plates, less than 1% blue colonies were detected.

To check for insert length, the inserts of 40 randomly picked white clones from *Eco*32I ligations were amplified by colony-PCR and separated in a 1% agarose gel. Thirty-seven of 40 selected clones (92.5%) from fraction I and 36 of 40 selected clones (90%) from fraction II contained an insert, demonstrating the effective blue/white screening of an *Eco*32I blunt-ended vector (Figure 1). The average insert length was 292 and 545 bp for fractions I and II, respectively. Similar insert lengths were found for *Sph*I clones (313 and 525 bp, respectively), demonstrating that the inherent T4 DNA polymerase exonuclease activity did not affect the concatemer length. The rather short insert size was due to a steep decline of long concatemers in the fractions. Two of the clones without insert (marked by a filled square in Figure 1) were due to vector religation; the other ones yielded likely no PCR products because of corrupted/missing primer site(s), as a subsequent analysis revealed (data not shown). Taken together, the total number of insert containing clones that can be obtained amounts to more than 330 000 for fraction I and 107 000 for fraction II after blunt-end ligation (Table 1). Moreover, greater than or equal to 90% of them contain an insert. This exceeds for each ligation by far the clone numbers ($< 3 \times 10^3$) typically analyzed during a SAGE experiment.

To prove finally that the treatment with T4 DNA polymerase did not affect the quality of the concatemers, several clones were sequenced. Again, no difference between polished and unpolished concatemers was observed (data not shown).

These data demonstrate that the conversion of the 3′-protruding termini of SAGE concatemers into blunt ends represents an efficient and robust cloning strategy. The additional step of blunt-ending takes less than 1 h and requires only T4 DNA polymerase. No further purification step is required because the reaction mixture is directly used for ligation, and the protocol can be used for every vector containing an *Eco*RV/*Eco*32I restriction endonuclease recognition site. This avoids problems associated with the often used pZERO vector (Invitrogen, Carlsbad, CA, USA), which was sometimes shown to be unstable, and eliminates the need of the otherwise rarely used antibiotic zeomycin. In conclusion, the advantages mentioned above make our protocol a versatile, efficient, and inexpensive technique that should prove very helpful for the construction of SAGE libraries.

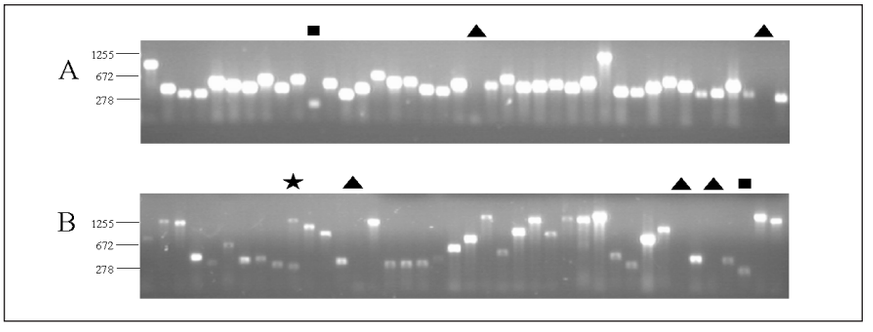

Figure 1. Colony-PCR amplification of SAGE library inserts cloned by blunt-end ligation with *Eco*32I-cut vector. Colony-PCR was performed by using M13 forward and reverse primers in a 50-µL reaction and the following conditions: 94°C for 30 s, 60°C for 30 s, and 72°C for 2 min for 35 cycles. Ten microliters of each reaction were analyzed in a 1% agarose gel. (A) Insert amplification of gel-purified concatemers of 300–500 bp (fraction I) in length. (B) Insert amplification of gel-purified concatemers of 500–1000 bp (fraction II) in length. Differences in the intensity of the insert bands were due to the applied technique (i.e., colony-PCR). The asterisk marks a colony-PCR in which two independent clones had been accidentally amplified; filled squares mark religated clones without an insert; filled triangles indicate clones where the PCR yielded no amplification product, probably because of missing primer binding site(s).

REFERENCES

1. **Velculescu, V.E., L. Zhang, B. Vogelstein, and K.W. Kinzler.** 1995. Serial analysis of gene expression. Science *270*:484-487.
2. **Jasper, H., V. Benes, A. Atzberger, S. Sauer, W. Ansorge, and D. Bohmann.** 2002. A genomic switch at the transition from cell proliferation to terminal differentiation in the Drosophila eye. Dev. Cell *3*:511-521.
3. **Boon, K., E.C. Osorio, S.F. Greenhut, C.F. Schaefer, J. Shoemaker, K. Polyak, P.J. Morin, K.H. Buetow, et al.** 2002. An anatomy of normal and malignant gene expression. Proc. Natl. Acad. Sci. USA *99*:11287-11292.
4. **Liang, P.** 2002. SAGE Genie: a suite with panoramic view of gene expression. Proc. Natl. Acad. Sci. USA *99*:11547-11548.
5. **Saha, S., A.B. Sparks, C. Rago, V. Akmaev, C.J. Wang, B. Vogelstein, K.W. Kinzler, and V.E. Velculescu.** 2002. Using the transcriptome to annotate the genome. Nat. Biotechnol. *20*:508-512.
6. **Datson, N.A., J. Perk-de Jong, M.P. van den Berg, E.R. de Kloet, and E. Vreugdenhil.** 1999. MicroSAGE: a modified procedure for serial analysis of gene expression in limited amounts of tissue. Nucleic Acids Res. *27*:1300-1307.
7. **Angelastro, J.M., L.P Klimaschewski, and O.V. Vitolo.** 2000. Improved NlaIII digestion of PAGE-purified 102 bp ditags by addition of a single purification step in both the SAGE and microSAGE protocols. Nucleic Acids Res. *28*:e62.
8. **Lee, S., J.J. Chen, G.L. Zhou, and S.M. Wang.** 2001. Generation of high-quantity and quality tag/ditag cDNAs for SAGE analysis. BioTechniques *31*:348-354.
9. **Peters, D.G., A.B. Kassam, H. Yonas, E.H. O'Hare, R.E. Ferrell, and A.M. Brufsky.** 1999. Comprehensive transcript analysis in small quantities of mRNA by SAGE-Lite. Nucleic Acids Res. *27*:e39.
10. **Powell, J.** 1998. Enhanced concatemer cloning—a modification to the SAGE (serial analysis of gene expression) technique. Nucleic Acids Res. *26*:3445-3446.
11. **Kenzelmann, M. and K. Muehlemann.**

1999. Substantially enhanced cloning efficiency of SAGE (serial analysis of gene expression) by adding a heating step to the original protocol. Nucleic Acids Res. *27*:917-918.
12. **Sambrook, J. and D.W. Russell.** 2001. Molecular Cloning: A Laboratory Manual. CSH Laboratory Press, Cold Spring Harbor, NY.
13. **Koesters, R. and M.V. Doeberitz.** 2002. Improved blunt-end cloning by replacing *Eco*RV with *Eco*32I. BioTechniques *32*:1244-1246.

This research was supported by DFG grant no. DFG No428/1-1 to H.G.N. and E.F. Address correspondence to Dr. H.G. Nothwang, Abt. Tierphysiologie, Universität Kaiserslautern, Postfach 3049, D-67653 Kaiserslautern, Germany. e-mail: nothwang@rhrk.uni-kl.de

Received 18 December 2002; accepted 28 January 2003.

Alexander Koehl, Eckhard Friauf, and Hans Gerd Nothwang
Universität Kaiserslautern
Kaiserslautern, Germany